目　录

第一章　矩　　阵

矩阵的概念起源于十八世纪，它是由解线性代数方程组以及将二次曲面(曲线)方程化为标准形的需要而逐步形成的。1856～1857 年，Cayley 首先提出了"矩阵"这一术语，一直沿用至今。本章引入矩阵及其运算，讨论它们的基本性质，包括有广泛应用的非异阵；介绍常用的矩阵分块方法、矩阵的初等变换以及运用标准单位向量的一些基本技巧。

§1　和 号 Σ

一、单重和号

为简便起见，将 n 个数 $a_1, a_2, \cdots, a_n$ 之和 $a_1+a_2+\cdots+a_n$ 记作 $\sum\limits_{i=1}^{n} a_i$，其中 a_i 称为一般项，i 是它的足标；Σ 称为和号，Σ 上、下的数字表示求和时足标 i 的取值范围。例如

$$1^2+3^2+\cdots+(2n-1)^2=\sum_{i=1}^{n}(2i-1)^2.$$

利用数的运算规则，容易验证下列等式成立：

$$\sum_{i=1}^{n}(a_i+b_i)=\sum_{i=1}^{n}a_i+\sum_{i=1}^{n}b_i; \tag{1}$$

$$\sum_{i=1}^{n}a_i=\sum_{i=1}^{l}a_i+\sum_{i=l+1}^{n}a_i \qquad (1\leqslant l<n)①; \tag{2}$$

$$\sum_{i=1}^{n}ka_i=k\sum_{i=1}^{n}a_i \qquad (k\text{ 为任一个与 } i \text{ 无关的数}). \tag{3}$$

---

① 由上下文不难看出，这里的 l 是一个适合条件 $1\leqslant l<n$ 的整数。今后，对类似的情形将不再附加说明。

在计算或化简和式时，应用这些等式将带来方便。

例 1 已知 $\sum_{i=1}^{n} a_i=\frac{1}{2}(n^2+n)$，计算 $\sum_{i=1}^{n}(a_i+1)$。

【解】 $\sum_{i=1}^{n}(a_i+1)=\sum_{i=1}^{n}a_i+\sum_{i=1}^{n}1=\frac{1}{2}(n^2+n)$

$$+\underbrace{(1+1+\cdots+1)}_{n\text{ 个}}=\frac{1}{2}n(n+3).$$

值得注意的是，$\sum_{i=1}^{n}a_i$ 与 $\sum_{j=1}^{n}a_j$ 同样代表 $a_1+a_2+\cdots+a_n$，即

$$\sum_{i=1}^{n}a_i=\sum_{j=1}^{n}a_j.$$

所以，只要不发生混淆，用什么字母作求和足标都可以。

对于同一个和，用和号缩写时，可以有几种表达式。例如

$$a_1+a_2+\cdots+a_n=\sum_{i=1}^{n}a_i=\sum_{i=0}^{n-1}a_{i+1},$$

这两种表达式中，一般项的记法不同，因而求和足标的取值范围也随之而不同。究竟取哪一种表达式，要看运算的需要而定。

二、双重和号

对带有两个足标的 mn 个数的和

$$S=a_{11}+a_{12}+\cdots+a_{1n}+a_{21}+a_{22}+\cdots+a_{2n}$$
$$+\cdots+a_{m1}+a_{m2}+\cdots+a_{mn},$$

用和号缩写时，可以先作各行（横排）中数的和，再作总和，得到

$$S=\sum_{j=1}^{n}a_{1j}+\sum_{j=1}^{n}a_{2j}+\cdots+\sum_{j=1}^{n}a_{mj}=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}\right); \tag{4}$$

也可先作各列（竖排）中数的和，再作总和，得到

$$S=\sum_{i=1}^{m}a_{i1}+\sum_{i=1}^{m}a_{i2}+\cdots+\sum_{i=1}^{m}a_{in}=\sum_{j=1}^{n}\left(\sum_{i=1}^{m}a_{ij}\right); \tag{5}$$

约定

$$\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}\equiv\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}\right),$$

由(4)、(5)两式可得

$$\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{m}a_{ij}. \tag{6}$$

这表明，在和式 $\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}$ 中，和号 $\sum_{i=1}^{m}$ 与 $\sum_{j=1}^{n}$ 的次序可以交换。

特别地，当 $m=n$ 时，有

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{n}a_{ij},\tag{7}$$

由此还可推出常用的等式

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ji}.\tag{8}$$

通常，将 $\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}$ 简记为 $\sum_{i,j=1}^{n}a_{ij}$。

例 2 证明：(i) $\sum_{i,j=1}^{n}x_ix_j=\left(\sum_{i=1}^{n}x_i\right)^2$；

(ii) $\sum_{j=1}^{n}\sum_{k=1}^{j}C_{kj}+\sum_{j=1}^{n}\sum_{k=j}^{n-1}C_{k+1,j}=\sum_{i,j=1}^{n}C_{ij}$。

【证明】 (i)

$$\begin{aligned}\sum_{i,j=1}^{n}x_ix_j&=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}x_ix_j\right)=\sum_{i=1}^{n}\left(x_i\sum_{j=1}^{n}x_j\right)\\&=\left(\sum_{j=1}^{n}x_j\right)\cdot\left(\sum_{i=1}^{n}x_i\right)\\&=\left(\sum_{i=1}^{n}x_i\right)^2;\end{aligned}$$

(ii)

$$\begin{aligned}\sum_{j=1}^{n}\sum_{k=1}^{j}C_{kj}+\sum_{j=1}^{n}\sum_{k=j}^{n-1}C_{k+1,j}&=\sum_{j=1}^{n}\sum_{k=1}^{j}C_{kj}+\sum_{j=1}^{n}\sum_{k=j+1}^{n}C_{kj}\\&=\sum_{j=1}^{n}\sum_{k=1}^{n}C_{kj}=\sum_{i,j=1}^{n}C_{ij}.\end{aligned}$$

最后指出，和号 Σ 作用的对象不限于数。譬如，对三维实向量，(1)～(3)式仍然成立(此时，(3) 式应理解为向量的数乘)。推而广之，对"加法"具有交换律与结合律、"数乘"对"加法"具有分配律的一切运算对象，相应的结论也成立。

习　题

1. 证明下列等式：

(i) $\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{j=0}^{n-1}x_{j+1}\right)=\left(\sum_{k=-1}^{n-2}x_{k+2}\right)^2$；

(ii) $\sum_{k=1}^{n}(2k+1)^2-\sum_{j=0}^{n-1}(2j+1)^2=4n(n+1)$，

(iii) $\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki}=\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ki}b_{ik}$.

2. 以 $\bar{Z}$ 代表复数 Z 的共轭复数。求证，对复数 y_i、$z_i(i=1,2,\cdots,n)$，下列等式成立

(i) $\overline{\sum_{i=1}^{n}(y_i+z_i)}=\sum_{i=1}^{n}\bar{y}_i+\sum_{i=1}^{n}\bar{z}_i$;

(ii) $\overline{\sum_{i=1}^{n}y_iz_i}=\sum_{i=1}^{n}\bar{y}_i\bar{z}_i$.

3. 化简下列各式：

(i) $\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-x_j)$;

(ii) $\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-x_j)^2+\sum_{j=1}^{n}\sum_{i=1}^{n}(x_j+x_i)^2$.

4. 记 $c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}(i,j=1,2,\cdots,n)$，若 $\sum_{i=1}^{n}a_{ik}=1(k=1,2,\cdots,n)$，求证

$$\sum_{i=1}^{n}(c_{ij}-b_{ij})=0(j=1,2,\cdots,n).$$

5. 证明：$\sum_{i=1}^{n}\sum_{k=1}^{i-1}a_{ik}=\sum_{i=1}^{n}\sum_{k=i+1}^{n}a_{ki}$.

6. 记 $\sum_{1\leqslant i\leqslant j\leqslant n}a_{ij}=\sum_{i=1}^{n}\sum_{j=i}^{n}a_{ij}$，

求证 $\sum_{1\leqslant i\leqslant j\leqslant n}a_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{j}a_{ij}$.

7. 记 $\sum_{1\leqslant i<j\leqslant n}a_{ij}=\sum_{i=1}^{n}\sum_{j=i+1}^{n}a_{ij}$，

求证 $\sum_{1\leqslant i<j\leqslant n}a_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{j-1}a_{ij}$.

8. 证明：$\sum_{i=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}a_{ik}=\sum_{k=1}^{n}\sum_{\substack{i=1\\i\neq k}}^{n}a_{ik}=\sum_{i=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}a_{ki}$.

§2 矩阵的概念及运算

一、矩阵的概念

现实世界中，事物之间的联系通常以一组变量与另一组变量之间的关系来描写。"线性"关系是其中最简单、最常见的一种。例如，将平面直角坐标系绕坐标原点旋转 θ 角得到新的坐标系之后，

平面上任一点的新、旧坐标(x, y)与(x', y')之间有如下的线性关系

$$x' = \cos\theta \cdot x + \sin\theta \cdot y$$
$$y' = -\sin\theta \cdot x + \cos\theta \cdot y$$

此外，技术科学及经济数学中的某些问题也往往归结为讨论一组变量$x_1, x_2, \cdots, x_n$与另一组变量$y_1, y_2, \cdots, y_n$之间的线性关系式：

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n} \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn} \end{cases} \tag{1}$$

其中，$a_{ij}(i=1, 2, \cdots, m; j=1, 2, \cdots, n)$是$mn$个常数。

若将(1)式中常数按其在(1)式中的相对位置排成矩形阵列

$$\begin{matrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots\cdots & \cdots\cdots \\ a_{m1} & a_{m2}\cdots a_{mn}, \end{matrix} \tag{2}$$

则(2)完全可以简单明确地代表线性关系式(1)。后面的几章将表明，要研究与线性关系式(1)有关的各种问题，必须先对矩形阵列(2)以及它们之间的“运算”作相应的讨论。

定义 由mn个数排成的m行(横的)n列(纵的)的矩形阵列(2)称为一个m行n列**矩阵**或 **$m \times n$ 阵**。

为了强调(2)是一个整体和避免今后运算中发生混淆，在(2)的两边添加一对圆括号。矩阵以大写的拉丁字母记之。通常一个$m \times n$阵可表示为

$$A = \begin{pmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots\cdots & \cdots\cdots \\ a_{m1} & a_{m2}\cdots a_{mn} \end{pmatrix}, \tag{3}$$

数$a_{ij}(i=1, 2, \cdots, m; j=1, 2, \cdots n)$称为矩阵$A$的元素。行足标$i$、列足标$j$指出$a_{ij}$是$A$的第$i$行第$j$列的元素，或简称为$A$的

(i,j)元素.(3)式亦可简记为 $A=(a_{ij})_{m\times n}$.矩阵的元素之间、元素的行足标与列足标之间,必要时可插入逗号,如$(2,b)$,$a_{p+q,r+s}$等。

当 $m=1$ 时,A 为 $1\times n$ 阵,也称之为**行矩阵**或 **n 维行向量**;当 $n=1$ 时,A 为 $m\times 1$ 阵,也称之为**列矩阵**或**m维列向量**. 行(列)向量中的元素也称为**分量**.

当 $m=n$ 时,A 为 $n\times n$ 阵,也称之为 **n 阶(方)阵**. 特别,当 $m=n=1$ 时,A 为 1×1 阵.常将(a)记为 a。

二、矩阵的运算

先引进相等的定义。

定义 设 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{l\times k}$, 称A与B相等(记为 $A=B$), 如果 $m=l$, $n=k$ 且 $a_{ij}=b_{ij}$ $(i=1,2,\cdots,m;\ j=1,2,\cdots,n)$。

1. *加法*

定义 设 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{m\times n}$, 称 $C=(a_{ij}+b_{ij})_{m\times n}$ 为 A 与 B 相加所得的和,记为 $C=A+B$。

例如

$$\begin{pmatrix} 2 & -1 & 1 \\ 3 & 2 & \frac{1}{2} \end{pmatrix}+\begin{pmatrix} 1 & 2 & \sqrt{-1} \\ -1 & \sqrt{3} & -1 \end{pmatrix}$$

$$=\begin{pmatrix} 3 & 1 & 1+\sqrt{-1} \\ 2 & 2+\sqrt{3} & -\frac{1}{2} \end{pmatrix}.$$

若 $A=(a_{ij})_{m\times n}$, 则称$(-a_{ij})_{m\times n}$ 为A的负矩阵, 记为 $-A$。

对 $m\times n$ 阵 A 与 B, 规定减法为

$$A-B=A+(-B).$$

元素全是零的矩阵称为零矩阵,以 O 记之。

不难验证,对矩阵的加法,有下列运算规则:

(i) $A+B=B+A$; (ii) $(A+B)+C=A+(B+C)$;

(iii) $A+O=A$; (iv) $A+(-A)=O$.

今证(i): 设 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{m\times n}$, 则 $A+B$ 为 $m\times n$ 阵,其

(i, j)元素为 $a_{ij}+b_{ij}$；$B+A$ 亦为 $m\times n$ 阵，其(i, j)元素为 $b_{ij}+a_{ij}$。因两数相加可交换，故 $a_{ij}+b_{ij}=b_{ij}+a_{ij}(i=1, 2, \cdots, m; j=1, 2, \cdots, n)$。据矩阵相等的定义，可得 $A+B=B+A$。类似地，可证明(ii)～(iv)。

上述运算规则 (ii)，使得我们可以这样来规定多个矩阵的加法：

$$A_1+A_2+\cdots+A_{s-1}+A_s=(A_1+A_2+\cdots+A_{s-1})+A_s \quad (s\geqslant 3).$$

2．数乘

定义 设 $A=(a_{ij})_{m\times n}$，k 为一个数，称 $C=(ka_{ij})_{m\times n}$ 为 k 与 A 相乘(数乘)所得的积，记为 $C=kA$。

约定 $Ak=(a_{ij}k)_{m\times n}$，则显然有 $kA=Ak$，故也称 Ak 为 k 与 A 数乘所得的积。

不难验证，对矩阵的数乘，有下列运算规则：

(i) $k(A+B)=kA+kB$； (ii) $(k+l)A=kA+lA$；

(iii) $(kl)A=k(lA)$； (iv) $1\cdot A=A$。

3．乘法

矩阵乘法是最主要的矩阵运算。先简述一下它的由来。考虑由线性关系式

$$z_i=\sum_{k=1}^{n} a_{ik}y_k \qquad (i=1, 2, \cdots, m) \tag{4}$$

$$y_k=\sum_{j=1}^{l} b_{kj}x_j \qquad (k=1, 2, \cdots, n) \tag{5}$$

出发，将 $z_1, z_2, \cdots, z_m$ 表示为 $x_1, x_2, \cdots, x_l$ 的线性关系式。将(5)代入(4)得

$$z_i=\sum_{k=1}^{n} a_{ik}\left(\sum_{j=1}^{l} b_{kj}x_j\right)=\sum_{j=1}^{l}\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)x_j \quad (i=1, 2, \cdots, m),$$

记

$$c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj} \qquad (i=1, 2, \cdots, m; j=1, 2, \cdots, l), \tag{6}$$

便得到 $z_1, z_2, \cdots, z_m$ 与 $x_1, x_2, \cdots, x_l$ 之间的线性关系式

$$z_i = \sum_{j=1}^{l} c_{ij} x_j \qquad (i=1, 2, \cdots, m). \tag{7}$$

若分别以 $A=(a_{ij})_{m\times n}$、$B=(b_{ij})_{n\times l}$、$C=(c_{ij})_{m\times l}$ 代表线性关系式(4)、(5)、(7)中的系数阵列，由(6)式可见，矩阵C的 (i, j) 元素恰好是矩阵A的第 i 行与矩阵B的第 j 列的对应元素乘积的和。

定义 设 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{n\times l}$, 令

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad (i=1, 2, \cdots, m; j=1, 2, \cdots, l),$$

称 $C=(c_{ij})_{m\times l}$ 为A与B相乘所得的积，记为 $C=AB$。

矩阵A与B可乘（也称 AB 有意义）的前提是A的列数等于B的行数。乘积矩阵 AB 的行数与列数分别等于A的行数与B的列数，其(i, j) 元素是A的第 i 行与B的第 j 列的对应元素乘积的和。

$$AB = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \vdots & b_{1j} & \vdots \\ \vdots & b_{2j} & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & b_{nj} & \vdots \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & c_{ij} & \cdots \\ & \vdots & \end{pmatrix} = C.$$

例如，设

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & -1 \\ -1 & 0 & -1 & 0 \end{pmatrix},$$

则

$$AB = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 3 & 1 & 4 & 1 \end{pmatrix}.$$

又如

$$(x_1, x_2, \cdots, x_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} x_i^2.$$

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}(b_1, b_2, \cdots, b_l) = \begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_l \\ a_2b_1 & a_2b_2 & \cdots & a_2b_l \\ \cdots & \cdots & \cdots & \cdots \\ a_mb_1 & a_mb_2 & \cdots & a_mb_l \end{pmatrix}.$$

若记

$$A = (a_{ij})_{m\times n},\quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},\quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix},$$

利用矩阵乘法，可将线性关系式(1)表示为

$$Y = AX.$$

不难验证，对于矩阵的乘法，有下列运算规则：

(i) $(AB)C = A(BC)$；

(ii) $(A+B)C = AC + BC$；$G(A+B) = GA + GB$；

(iii) $k(AB) = (kA)B = A(kB)$ （k 为任一个数）。

今证(i)：设 $A = (a_{ij})_{m\times n}$，$B = (b_{ij})_{n\times p}$，$C = (c_{ij})_{p\times q}$，令

$$AB = U = (u_{ij})_{m\times p},\ BC = V = (v_{ij})_{n\times q},$$

则$(AB)C$ 与 $A(BC)$都是 $m\times q$ 阵。又$(AB)C$ 的(i, j)元素为

$$\sum_{k=1}^{p} u_{ik}c_{kj} = \sum_{k=1}^{p}\left(\sum_{l=1}^{n} a_{il}b_{lk}\right)c_{kj} = \sum_{l=1}^{n} a_{il}\left(\sum_{k=1}^{p} b_{lk}c_{kj}\right)$$
$$= \sum_{l=1}^{n} a_{il}v_{lj},$$

它恰好等于 $A(BC)$的(i, j)元素$(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$，于是得到$(AB)C = A(BC)$。类似地，可证明(ii)与(iii)。

上述运算规则 (i)，使得我们可以这样来规定多个矩阵的乘法：

$$A_1A_2\cdots A_{s-1}A_s = (A_1A_2\cdots A_{s-1})A_s \quad (s\geqslant 3).$$

值得注意的是，对于矩阵乘法，一般说来，不满足交换律，即"$AB = BA$" 未必对所有的 A 和 B 都成立。这是因为，若 AB 有意

义，BA 未必有意义；即使 BA 也有意义，此时 AB 与 BA 均为方阵，但它们的阶数未必相同；即使 AB 与 BA 为同阶方阵，也未必有 $AB=BA$。如 $A=\begin{pmatrix}0&0\\0&1\end{pmatrix}$，$B=\begin{pmatrix}0&1\\0&0\end{pmatrix}$，则 $AB=\begin{pmatrix}0&0\\0&0\end{pmatrix}$，而 $BA=\begin{pmatrix}0&1\\0&0\end{pmatrix}$。

当 $AB=BA$ 时，称 A 与 B 可交换。由上面的讨论易知，A 与 B 可交换的必要条件是 A 与 B 为同阶方阵。

为了区别矩阵相乘的次序，称 AB 为 A 左乘 B 的积或 B 右乘 A 的积。

上例还说明，即使 $A\neq O$，$B\neq O$，却可能有 $AB=O$。这是矩阵乘法的又一特点，称为**有零因子**。由此可知，对于矩阵乘法来说，一般不满足"消去律"，即若 $AB=AC$，$A\neq O$，未必有 $B=C$。对此，读者不难举出反例。

三、矩阵的转置

定义 设 $A=(a_{ij})_{m\times n}$，令

$$b_{ij}=a_{ji}(i=1,2,\cdots n;j=1,2,\cdots,m),$$

称 $B=(b_{ij})_{n\times m}$ 为 A 取转置后得到的转置矩阵，记为 $B=A'$。

由定义，若

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},$$

则

$$A'=\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{m1}\\a_{12}&a_{22}&\cdots&a_{m2}\\\cdots&\cdots&\cdots&\cdots\\a_{1n}&a_{2n}&\cdots&a_{mn}\end{pmatrix}.$$

可见，A 的第 i 行元素依次是 A' 的第 i 列元素$(i=1,2,\cdots,m)$；

A 的第 j 列元素依次是 A' 的第 j 行元素($j=1, 2, \cdots, n$)。简言之,将矩阵转置,即依次将其行改写成列.

不难验证,对于矩阵的转置,有下列运算规则:

(i) $(A')'=A$;

(ii) $(A+B)'=A'+B'$;

(iii) $(kA)'=kA'$(k 为任一个数);

(iv) $(AB)'=B'A'$。

今证(iv):设 $A=(a_{ij})_{m\times n}, B=(b_{ij})_{n\times p}, A'=(\tilde{a}_{ij})_{n\times m}, B'=(\tilde{b}_{ij})_{p\times n}$,则$(AB)'$ 与 $B'A'$ 均为 $p\times m$ 阵。由定义,$(AB)'$ 的(i, j)元素等于 AB 的(j, i)元素,即应为

$$\sum_{k=1}^{n} a_{jk}b_{ki} \qquad (i=1, 2, \cdots, p; j=1, 2, \cdots, m).$$

而 $B'A'$ 的(i, j)元素为

$$\sum_{k=1}^{n} \tilde{b}_{ik}\tilde{a}_{kj}=\sum_{k=1}^{n} b_{ki}a_{jk}=\sum_{k=1}^{n} a_{jk}b_{ki}$$

$$(i=1, 2, \cdots, p; j=1, 2, \cdots m).$$

于是得到$(AB)'=B'A'$。类似地,可证明(i)~(iii)。

四、方阵

方阵是矩阵中重要的一类。自然,矩阵的所有运算及性质都适用于方阵。

1. 对角阵与单位阵

定义 称 n 阶方阵

$$A=\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}^{①}$$

为 **n 阶对角阵**,简记为 $A=[a_1, a_2, \cdots, a_n]$。特别,当对角元 a_1,

① 约定,矩阵中空白处的元素为零。

$a_2, \cdots, a_n$ 全为 1 时，称它为 n 阶**单位阵**，记为

$$I_n = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} = [1, 1, \cdots, 1].$$

记号

$$\delta_{ij} = \begin{cases} 1 & \text{当 } i = j; \\ 0 & \text{当 } i \neq j. \end{cases}$$

通常称为 Kronecker 符号 δ。利用它，对角阵 $A = [a_1, a_2, \cdots, a_n]$ 的 (i, j) 元素可以表示为 $a_i\delta_{ij}$ 或 $a_j\delta_{ij}$。特别地，$I_n = (\delta_{ij})_{n\times n}$。

对任一 n 阶方阵 A，恒有 $I_nA = AI_n = A$。这是单位阵的重要性质。它表明，n 阶单位阵在 n 阶方阵中的地位类似于 1 在数中的地位。有时，无须指明单位阵的阶数，也将 I_n 记为 I。

称 kI(k 为任意一个数) 为**纯量矩阵**。显然，纯量矩阵是对角阵。对任一方阵 A，恒有

$$(kI)A = A(kI) = k(AI) = kA,$$

上式表明，在与方阵相乘时，纯量矩阵 kI 与数 k 具有同样的作用。

2. 方阵的幂

定义 设 A 为方阵，规定

$A^0 = I; A^k = A^{k-1}A$ （k 为一正整数），

称 A^k 为方阵 A 的 k 次幂。

不难验证，对于方阵的幂，有下列“指数公式”：

(i) $A^kA^l = A^{k+l}$； (ii) $(A^k)^l = A^{kl}$；

(iii) 若 A、B 为同阶方阵，且 $AB = BA$，则

$$(AB)^k = A^kB^k,$$

其中，k、l 为非负整数。

习　题

1. 计算

(i) $\begin{pmatrix} 1+\sqrt{-1} & -1 \\ 2 & 1-\sqrt{-1} \\ 0 & \sqrt{-1} \end{pmatrix} \begin{pmatrix} 1-\sqrt{-1} & \\ & -1 \end{pmatrix}$;

(ii) $(x,y,1)\begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$;

(iii) $\begin{pmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \vdots \\ 1 & a_n \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}$;

(iv) $\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} 1 & b_1 \\ 1 & b_2 \\ \vdots & \vdots \\ 1 & b_n \end{pmatrix}$;

(v) $x'Ax$,其中 $x'=(x_1,x_2,\cdots,x_n)$,$A=(a_{ij})_{n\times n}$。

2. 证明:

(i) $(A\pm B)C=AC\pm BC$; (ii) $G(A\pm B)=GA\pm GB$。

3. 求所有与 A 可交换的矩阵:

(i) $A=\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$; (ii) $A=\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$。

4. 设实方阵 $A=B+C$,且 $B=B'$,$-C=C'$,求证 $AA'=A'A \Longleftrightarrow BC=CB$。

5. 设 n 阶方阵 $M=I_n-2\alpha\alpha'$,其中,α 为实的 n 维列向量,且 $\alpha'\alpha=1$(称此 M 为**实镜象阵**)。求证 $M=M'$,且 $MM'=I$。

6. 称实方阵 A 为**对称阵**,如果 $A=A'$;称实方阵 A 为**反对称阵**,如果 $A=-A'$。求证:

(i) 若 A、B 为同阶对称阵,则 AB 为对称阵 $\Longleftrightarrow AB=BA$。举例说明,两对称阵之积未必是对称阵;

(ii) 若 A 为对称阵,B 为同阶反对称阵,则 AB 为反对称阵 $\Longleftrightarrow AB=BA$。

7. 设 $D=[d_1,d_2,\cdots,d_n]$,其中,当 $i\neq j$ 时,$d_i\neq d_j$ $(i,j=1,2,\cdots,n)$。求证与 D 可交换的矩阵只能是对角阵。

8. 计算

(i) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{100}$，　(ii) $\begin{pmatrix} \frac{1}{3} & -\frac{2\sqrt{2}}{3} \\ -\frac{2\sqrt{2}}{3} & -\frac{1}{3} \end{pmatrix}^{50}$.

9. (i) 给出二阶方阵 $A\neq O$、$B\neq O$，使 $A^2+B^2=O$；(ii) 设 A、B 为同阶方阵，求证 $A^2-B^2=(A+B)(A-B)$、$(A+B)^2=A^2+2AB+B^2$ 成立的充分必要条件是 $AB=BA$。

10. 称方阵 A 为**幂等阵**，如果 $A^2=A$。

(i) 验证方阵 A 为幂等阵：

$$A=\begin{pmatrix} -26 & -18 & -27 \\ 21 & 15 & 21 \\ 12 & 8 & 13 \end{pmatrix};$$

(ii) 记 $A=\frac{1}{2}(I\pm B)$。求证：A 为幂等阵 $\Longleftrightarrow B^2=I$。

11. 称方阵 A 为**对合阵**，如果 $A^2=I$。

(i) 验证方阵 A 为对合阵，

$$A=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & -3 & 3 & -1 \end{pmatrix};$$

(ii) 设 A 为对合阵，$P=AD_1A$，$Q=AD_2A$，D_1、D_2 为对角阵。求证 $PQ=QP$。

12. 设 $A=(a_{ij})_{m\times n}$，以 $\overline{a_{ij}}$ 表示复数 a_{ij} 的共轭复数，称 $B=(\overline{a_{ij}})_{m\times n}$ 为 A **共轭矩阵**，记为 $B=\overline{A}$。验证下列运算规则：

(i) $\overline{(\overline{A})}=A$；　(ii) $\overline{A+B}=\overline{A}+\overline{B}$；　(iii) $\overline{kA}=\overline{k}\,\overline{A}$　(k 为任一复数)；　(iv) $\overline{AB}=\overline{A}\,\overline{B}$；(v) $\overline{A'}=\overline{A}'$。

§3　非异阵、逆阵

一、非异阵及其逆阵

在 §2 中，讨论了矩阵的加、减、乘运算。现在要问，矩阵有无"除"的运算？本节仅就方阵来讨论这一问题。

对于数a,只要$a\neq 0$,则必存在它的倒数$\frac{1}{a}$,且$a\cdot\frac{1}{a}=\frac{1}{a}\cdot a=1$;此时,任一个数$b$除以$a$,可以表示为$b\cdot\frac{1}{a}$。相仿地,要谈及方阵的"除",首先就要问,对于方阵A,若$A\neq O$,是否存在方阵B,能使$AB=BA=I$?回答并非是肯定的。如$A=\begin{pmatrix}0&0\\0&1\end{pmatrix}$,虽然$A\neq O$,但任一个二阶方阵$B$都不能使$AB=BA=I$。这是因为,$AB$的(1, 1)元素恒为零,而$I_2$的(1, 1)元素恒为1。

定义 称方阵A为**非异阵**,如果存在同阶方阵B,使得

$$AB=BA=I, \tag{1}$$

否则,称A为**奇异阵**。称适合等式(1)的方阵B为方阵A的**逆阵**。非异阵也称为**可逆阵**。

命题1 如果A有逆阵,则逆阵是唯一的。

【证明】 设B_1、B_2都适合等式(1):

$$AB_1=B_1A=I;AB_2=B_2A=I,$$

则

$$B_1=IB_1=(B_2A)B_1=B_2(AB_1)=B_2I=B_2.$$ 证毕。

因逆阵是唯一的,故将A的逆阵记为A^{-1}。显然,若A为非异阵,则A^{-1}总存在,且$AA^{-1}=A^{-1}A=I$。

二、逆阵的运算

命题2 若A为非异阵,则A^{-1}、kA(k为任一个非零的数)、A'以及$\overline{A}$也都是非异阵,且

$$(A^{-1})^{-1}=A;\ (kA)^{-1}=\frac{1}{k}A^{-1}\qquad(k\neq 0);$$

$$(A')^{-1}=(A^{-1})';(\overline{A})^{-1}=\overline{A^{-1}}.$$

【证明】 仅对A^{-1}、A'证之,其余从略。因A为非异阵,便有$AA^{-1}=A^{-1}A=I$。据定义,此式表明A^{-1}是非异阵,且A是它的逆阵。又逆阵是唯一的,故得$(A^{-1})^{-1}=A$。将$AA^{-1}=A^{-1}A=I$转置,得到$(A^{-1})'A'=A'(A^{-1})'=I$.同理可得$A'$是非异阵,且

$(A')^{-1}=(A^{-1})'$. 证毕。

命题 3 若 A、B 是同阶非异阵，则AB 也是非异阵，且$(AB)^{-1}=B^{-1}A^{-1}$。

【证明】 因 A、B 是非异阵，故 A^{-1}、B^{-1} 存在，又

$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AIA^{-1}=AA^{-1}=I$；

$(B^{-1}A^{-1})(AB)=B^{-1}(A^{-1}A)B=B^{-1}IB=B^{-1}B=I$，

由定义可知，AB 是非异阵，且$(AB)^{-1}=B^{-1}A^{-1}$。 证毕。

由命题 3，应用归纳法，显然可得如下的推论。

推论 若 $A_1, A_2, \cdots, A_s(s\geqslant 2)$是同阶非异阵，则 $A_1A_2\cdots A_s$ 也是非异阵，且$(A_1A_2\cdots A_s)^{-1}=A_s^{-1}A_{s-1}^{-1}\cdots A_1^{-1}$。

以上两个命题的证明告诉我们，对于方阵 A，若能找到方阵 B，使得 $AB=BA=I$，则 $A^{-1}=B$，且A的非异性随之可得。

例 1 设A为幂等阵 ($A^2=A$)。求证 $I-2A$ 是非异阵，且 $(I-2A)^{-1}=I-2A$。

【证明】 因 $A^2=A$，于是$(I-2A)(I-2A)=I-4A+4A^2=I$，这表明 $I-2A$ 是非异阵，且 $(I-2A)^{-1}=I-2A$。 证毕。

但是，一般说来，对给定的方阵 A，要找适合(1)式的方阵 B，并不都是容易的，有时甚至很困难。因此，用例 1 那样的作法还不能完全解决方阵的非异性及求逆阵的问题。在本章第五节以及第二章中，对此将作进一步的讨论。

最后，举一例说明，如何利用命题 3 的结果来证明方阵非异并求其逆阵。

例 2 设 A、B 以及 $A+B$ 都是非异阵。求证 $A^{-1}+B^{-1}$ 也是非异阵，并求其逆阵。

【证明】 将 $A^{-1}+B^{-1}$ 表示成已知的非异阵的乘积：

$$A^{-1}+B^{-1}=A^{-1}(I+AB^{-1})=A^{-1}(B+A)B^{-1}, \tag{2}$$

由命题 3 的推论可知 $A^{-1}+B^{-1}$ 为非异阵，且

$$(A^{-1}+B^{-1})^{-1}=B(A+B)^{-1}A.$$

类似地，$A^{-1}+B^{-1}$ 也可表示为

$$A^{-1}+B^{-1}=B^{-1}(BA^{-1}+I)=B^{-1}(B+A)A^{-1}, \tag{3}$$

从而有

$$(A^{-1}+B^{-1})^{-1}=A(A+B)^{-1}B.$$ 证毕。

(2) 式(或(3)式)演变中的主要一步,在于将 A^{-1}、B^{-1} 从和中提取出来作为乘积的因子,这一技巧在第二章中还将被运用。

习　题

1. 证明:有一列(行)元素全是零的方阵必定是奇异阵。

2. 证明:奇异阵与非异阵之积必定是奇异阵。

3. 对于方阵 A,若存在矩阵 $C\neq O$,使得 $AC=O$,求证 A 为奇异阵。

4. 设 A 为对合阵,且 $A\neq\pm I$,

(i) 求证 A 为非异阵,并求出 A^{-1};

(ii) 求证 $I\pm A$ 都是奇异阵。

5. 设 A、B、C 为同阶方阵,

(i) 若 A 为非异阵,求证由 $AB=AC$ 可推得 $B=C$;

(ii) 若 A 为奇异阵,举例说明(i)的结论不成立。

6. 称实方阵 A 为**实正规阵**,如果 $AA'=A'A$。设 A 为非异的实正规阵,求证 A^{-1} 也是实正规阵。

7. 称实方阵 A 为**正交阵**,如果 $AA'=A'A=I$。设 A、B 为同阶正交阵,求证

(i) $A^{-1}=A'$,且 A^{-1}、A' 均为正交阵;

(ii) AB 为正交阵;

(iii) 对任一正交阵 P,$P^{-1}AP$ 为正交阵;

(iv) 若 $A+B$ 为正交阵,则 $(A+B)^{-1}=A^{-1}+B^{-1}$。

8. 求证:若方阵 A 满足下述三个条件中的任意两个,则必满足第三个:(i) A 为正交阵;(ii) A 为对称阵;(iii) A 为实对合阵。

9. 设 A 为 $m\times n$ 阵,且 $A'A$ 是非异阵,令 $B=I_m-A(A'A)^{-1}A'$,求证 $B'=B=B^2$ (即 B 是对称、幂等阵)。

10. 设 A、B 为同阶非异阵。求证:

(i) $A^{-1}B$、AB^{-1}、BA^{-1}、$B^{-1}A$ 也是非异阵,并求它们的逆阵;

(ii) A^{-1} 与 B^{-1} 可交换 $\Longleftrightarrow$ A 与 B 可交换。

11. 设 A、B 为同阶非异实方阵。求证:$A'A=B'B$ $\Longleftrightarrow$ 存在同阶正交阵 Q,使得 $A=QB$。

12. 设 A、B 为 n 阶非异阵，且 $AB-I_n$ 也为非异阵，求证

(i) $A-B^{-1}$ 为非异阵；

(ii) $(A-B^{-1})^{-1}-A^{-1}$ 也为非异阵，并求其逆阵。

§4 分块矩阵、标准单位向量

一、分块矩阵的概念

将一矩阵用横直线和纵直线划分成若干块，就得到"分块矩阵"。如

$$A=\left(\begin{array}{cc|c}1&0&0\\0&1&0\\\hline -1&0&2\end{array}\right)$$

就是一个分块矩阵。若记

$$A_{11}=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad A_{12}=\begin{pmatrix}0\\0\end{pmatrix},\ A_{21}=(-1,\ 0),\ A_{22}=2,$$

则 A 可表示为

$$A=\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix}.$$

沿用矩阵中的名称，称 A 为一个 2×2 分块矩阵，其中元素 $A_{ij}(i,\ j=1,\ 2)$ 是一些小矩阵。

一般地，将一 $m\times n$ 阵 A 用横直线划分成 r 块，用纵直线划分成 s 块，就得到一个 $r\times s$ 分块矩阵，通常记为

$$A=\begin{matrix}\\m_1\\m_2\\\vdots\\m_r\end{matrix}\begin{matrix}\begin{matrix}n_1&n_2&\cdots&n_s\end{matrix}\\\begin{pmatrix}A_{11}&A_{12}&\cdots&A_{1s}\\A_{21}&A_{22}&\cdots&A_{2s}\\\cdots&\cdots&\cdots&\cdots\\A_{r1}&A_{r2}&\cdots&A_{rs}\end{pmatrix}\end{matrix}=(A_{ij})_{r\times s},$$

其中，元素 $A_{ij}(i=1,\ 2,\ \cdots,\ r;\ j=1,\ 2,\ \cdots,\ s)$ 是 $m_i\times n_j$ 阵，也称它们为 A 的子块。由分块方法知道，

$$\sum_{i=1}^{r} m_i = m,\ \sum_{j=1}^{s} n_j = n,$$

所以，只要不改变子块之间的相对位置，将 $(A_{ij})_{r\times s}$ 看作以数为元素的矩阵，它仍是原先的 $m\times n$ 阵 A。

同一矩阵可进行不同的分块而得到不同的分块矩阵。例如，上述三阶方阵 A 也可作如下的分块：

$$A=\left(\begin{array}{c|cc} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{array}\right),$$

得到一个 1×2 分块矩阵。

对矩阵进行分块的目的在于揭示矩阵中某些部分的特性及它们之间的相互关系，有时也为了简化矩阵的计算。因此，矩阵如何分块，要看需要而定。

二、分块矩阵的运算

设 A、B 均为 $m\times n$ 阵，对 A、B 作如下同样的分块：

$$A=\begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array}\begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \cdots & n_s \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix} \end{array},$$

$$B=\begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array}\begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \cdots & n_s \end{array} \\ \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ B_{r1} & B_{r2} & \cdots & B_{rs} \end{pmatrix} \end{array}$$

则容易得到

$$A=B \Longleftrightarrow A_{ij}=B_{ij} \qquad (i=1,2,\cdots,r;\ j=1,2,\cdots,s);$$

以及

$$A' = \begin{array}{c} \\ n_1 \\ n_2 \\ \vdots \\ n_s \end{array} \begin{array}{c} \begin{array}{cccc} m_1 & m_2 & \cdots & m_r \end{array} \\ \begin{pmatrix} A'_{11} & A'_{21} & \cdots & A'_{r1} \\ A'_{12} & A'_{22} & \cdots & A'_{r2} \\ \cdots & \cdots & \cdots & \cdots \\ A'_{1s} & A'_{2s} & \cdots & A'_{rs} \end{pmatrix} \end{array}.$$

对于分块矩阵的乘法，有如下重要结论。

定理 4.1 设 A 为 $m \times n$ 阵，B 为 $n \times l$ 阵，若对 A、B 作如下分块：

$$A = \begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \cdots & n_s \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix} \end{array},$$

$$B = \begin{array}{c} \\ n_1 \\ n_2 \\ \vdots \\ n_s \end{array} \begin{array}{c} \begin{array}{cccc} l_1 & l_2 & \cdots & l_t \end{array} \\ \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ B_{s1} & B_{s2} & \cdots & B_{st} \end{pmatrix} \end{array}, \tag{1}$$

则

$$AB = \begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} l_1 & l_2 & \cdots & l_t \end{array} \\ \begin{pmatrix} G_{11} & G_{12} & \cdots & G_{1t} \\ G_{21} & G_{22} & \cdots & G_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ G_{r1} & G_{r2} & \cdots & G_{rt} \end{pmatrix} \end{array},$$

其中

$$G_{ij} = \sum_{k=1}^{s} A_{ik} B_{kj} \qquad (i = 1, 2, \cdots, r; j = 1, 2, \cdots, t). \tag{2}$$

【证明】 记

$$G=\begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array}\begin{array}{c} \begin{array}{cccc} l_1 & l_2 & \cdots & l_t \end{array} \\ \begin{pmatrix} G_{11} & G_{12} & \cdots & G_{1t} \\ G_{21} & G_{22} & \cdots & G_{2t} \\ \cdots & \cdots & \cdots & \cdots \\ G_{r1} & G_{r2} & \cdots & G_{rt} \end{pmatrix} \end{array},$$

下面证明，将G看作以数为元素的矩阵，有 $G=AB$。

首先，AB 为 $m\times l$ 阵。基于(1)的分块方式及(2)式，G_{ij} 为 $m_i\times l_j$ 阵，且 $\sum\limits_{i=1}^{r}m_i=m$，$\sum\limits_{j=1}^{t}l_j=l$，故将$G$看作以数为元素的矩阵，$G$ 也是一个 $m\times l$ 阵。

其次，$m\times l$ 阵G的 (i,j) 元素 g_{ij} 必位于分块矩阵G的某一子块 G_{pq} 之中，不妨设它是 G_{pq} 的(i',j')元素，即有

$$\begin{aligned} &i=m_1+m_2+\cdots+m_{p-1}+i',\ \text{其中}\ 1\leqslant i'\leqslant m_p;\\ &j=l_1+l_2+\cdots+l_{q-1}+j',\ \text{其中}\ 1\leqslant j'\leqslant l_q. \end{aligned}\tag{3}$$

由(2)式，

$$G_{pq}=A_{p1}B_{1q}+A_{p2}B_{2q}+\cdots+A_{ps}B_{sq},$$

可知 G_{pq} 的(i',j')元素应是 $A_{p1}, A_{p2}, \cdots, A_{ps}$ 的第 i' 行与 B_{1q}, $B_{2q}, \cdots, B_{sq}$ 的第 j' 列的相应元素乘积的和。由(3)式可知，A_{pk} 的第 i' 行元素位于A中第 i 行，B_{kq} 的第 j' 列元素位于B中第 j 列($k=1, 2, \cdots, s$)，再注意到对 A、B 所作的分块，可得

$$\begin{aligned} g_{ij}&=\sum_{k=1}^{n_1}a_{ik}b_{kj}+\sum_{k=n_1+1}^{n_1+n_2}a_{ik}b_{kj}+\cdots+\sum_{k=n_1+n_2+\cdots+n_{s-1}+1}^{n_1+n_2+\cdots+n_s}a_{ik}b_{kj}\\ &=\sum_{k=1}^{n}a_{ik}b_{kj}, \end{aligned}$$

这表明，矩阵G的(i,j)元素恰等于矩阵 AB 的(i,j)元素。

基于以上两点，可得 $G=AB$。　　　　证毕。

定理 4.1 表明，当矩阵A与B可乘时，只要对它们作恰当的分块(如(1)式那样，要求对A的纵向分块法与对B的横向分块法完全一致)，则所得分块矩阵便可按矩阵乘法的规则相乘(将子块当作数那样对待)，其积恰好等于A与B的积 AB。

分块矩阵的乘法是最主要的分块矩阵运算，它将应用于本书

的各个章节。现举例说明之。

例 1　设方阵

$$P=\begin{array}{c} \\ r \\ n-r\end{array}\begin{array}{c}\begin{array}{cc} r & n-r\end{array}\\ \begin{pmatrix} A & B\\ O & C\end{pmatrix}\end{array}.$$

其中，A、C 均为非异阵。求证 P 为非异阵，并求 P 的逆阵。

【证明】　先推测一下 P 的逆阵(如果存在的话)形式。设 Q 为 P 的逆阵，将 Q 作如下分块：

$$Q=\begin{array}{c} \\ r \\ n-r\end{array}\begin{array}{c}\begin{array}{cc} r & n-r\end{array}\\ \begin{pmatrix} Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{pmatrix}\end{array},$$

由 $PQ=I$ 及分块矩阵的乘法可得：

$$\begin{pmatrix} AQ_{11}+BQ_{21} & AQ_{12}+BQ_{22}\\ CQ_{21} & CQ_{22}\end{pmatrix}=\begin{pmatrix} I_r & O\\ O & I_{n-r}\end{pmatrix},$$

据分块矩阵相等的有关结论，应有

$AQ_{11}+BQ_{21}=I_r$, $AQ_{12}+BQ_{22}=O$, $CQ_{21}=O$, $CQ_{22}=I_{n-r}$。因 C 非异，由第三、四式可得 $Q_{21}=O$, $Q_{22}=C^{-1}$；分别将它们代入第一、二式，且因 A 非异，可得 $Q_{11}=A^{-1}$; $Q_{12}=-A^{-1}BC^{-1}$。于是

$$Q=\begin{pmatrix} A^{-1} & -A^{-1}BC^{-1}\\ O & C^{-1}\end{pmatrix}.$$

容易验证，$PQ=QP=I$。据此可断定，P 为非异阵，且其逆即为 Q。　　证毕。

特别地，若 A、C 为非异阵，则

$$P=\begin{pmatrix} A & O\\ O & C\end{pmatrix}$$

亦为非异阵，且

$$P^{-1}=\begin{pmatrix} A^{-1} & O\\ O & C^{-1}\end{pmatrix}.$$

定义　称分块方阵

$$A=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}$$

为**分块对角阵**，其中 $A_i(i=1, 2, \cdots, s)$为方阵，并将它简记为$A=[A_1, A_2, \cdots, A_s]$。

设 $A=[A_1, A_2, \cdots, A_s]$，$B=[B_1, B_2, \cdots, B_s]$，其中 A_i 与 $B_i(i=1, 2, \cdots, s)$为同阶方阵，则容易得到

$$AB=[A_1B_1, A_2B_2, \cdots, A_sB_s].$$

若 $A_1, A_2, \cdots, A_s$ 均为非异阵，注意到例 1 中 $B=O$ 的特款并用归纳法，可得

$$A^{-1}=[A_1^{-1}, A_2^{-1}, \cdots, A_s^{-1}].$$

三、标准单位向量及其应用

定义 称 n 维列向量

$$e_j=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} j\text{行}$$

为(第 j 个) n 维**标准单位(列)向量**，其中 j 可取 1 至 n 中间任一个正整数。

容易验证

$$e_i'e_j=\delta_{ij} \qquad (i, j=1, 2, \cdots, n). \tag{4}$$

今后，若不指明标准单位向量的维数，则总认为它的维数满足可乘的条件。

命题 1 对任一 $m\times n$ 阵 A，成立

$$Ae_j=\alpha_j; \tag{5}$$

$$e_i'A=\tilde{\alpha}_i, \tag{6}$$

其中，α_j 为 A 的第 j 列 $(j=1, 2, \cdots, n)$，$\tilde{\alpha}_i$ 为 A 的第 i 行 $(i=1, 2, \cdots, m)$。

【证明】 将 A 按列分块为

$$A=(\alpha_1, \alpha_2, \cdots, \alpha_n),$$

则

$$Ae_j=(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}=\alpha_j \quad (j=1, 2, \cdots, n);$$

将 A 按行分块为

$$A=\begin{pmatrix}\tilde{\alpha}_1\\ \tilde{\alpha}_2\\ \vdots\\ \tilde{\alpha}_m\end{pmatrix},$$

类似地可得

$$e_i'A=\tilde{\alpha}_i \qquad (i=1, 2, \cdots, m).$$ 证毕。

利用标准单位向量，配以适当的矩阵分块，可使某些矩阵的运算简捷而明瞭。这一方法，今后将要反复地运用，现举几例说明之。

例 2 称 $P_{j_1j_2\cdots j_n}=(e_{j_1}, e_{j_2}, \cdots, e_{j_n})$ 为 n 阶**排列阵**，其中 $j_1j_2\cdots j_n$ 为 $1,2,\cdots,n$ 的任意一个排列。求证，n 阶排列阵 $P_{j_1j_2\cdots j_n}$ 是正交阵。

【证明】 设 $P_{j_1j_2\cdots j_n}=(e_{j_1}, e_{j_2}, \cdots, e_{j_n})$，其转置矩阵为

$$P'_{j_1j_2\cdots j_n}=\begin{pmatrix}e'_{j_1}\\ e'_{j_2}\\ \vdots\\ e'_{j_n}\end{pmatrix},$$

注意到(4)式，可知 $P'_{j_1j_2\cdots j_n}P_{j_1j_2\cdots j_n}$ 的(k, l)元素为

$$e'_{j_k}e_{j_l}=\delta_{j_kj_l}=\delta_{kl},$$

于是得到

$$P'_{j_1j_2\cdots j_n}P_{j_1j_2\cdots j_n}=I; \tag{7}$$

若将 $P_{j_1j_2\cdots j_n}$ 按行分块，可记为

$$P_{j_1j_2\cdots j_n}=\begin{pmatrix} e'_{i_1} \\ e'_{i_2} \\ \vdots \\ e'_{i_n} \end{pmatrix},$$

其中，$i_1i_2\cdots i_n$ 为 $1, 2, \cdots, n$ 的某一个排列。取转置得到

$$P'_{j_1j_2\cdots j_n}=(e_{i_1}, e_{i_2}, \cdots, e_{i_n}),$$

类似地，可以证明

$$P_{j_1j_2\cdots j_n}P'_{j_1j_2\cdots j_n}=I. \tag{8}$$

综合(7)、(8)两式，可知 $P_{j_1j_2\cdots j_n}$ 为正交阵。 **证毕.**

例 3 设 n 阶方阵

$$A=\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},$$

求证 $A^n=O$.

【证明】 注意到 $A=(0, e_1, e_2, \cdots, e_{n-1})$，故

$$A^2=A(0, e_1, e_2, \cdots, e_{n-1})=(0, Ae_1, Ae_2, \cdots, Ae_{n-1}),$$

利用(5)式，应有

$$Ae_1=0,\ Ae_2=e_1,\ \cdots,\ Ae_{n-1}=e_{n-2},$$

于是

$$A^2=(0, 0, e_1, \cdots, e_{n-2}).$$

依次类推，可得

$$A^k=(\overbrace{0, \cdots, 0}^{k\text{ 个}}, e_1, e_2, \cdots, e_{n-k}) \quad (k=1, 2, \cdots, n).$$

当 $k=n$ 时,即为 $A^n=O$。 证毕。

例 4 设 A 为 $m\times n$ 阵。证明:若对任一 n 维列向量 x,恒有 $Ax=0$,则 $A=O$。

【证明】 既然对任一 n 维列向量 x 恒有 $Ax=0$,那末,特别取 $x=e_j(j=1, 2, \cdots, n)$,也应有

$$Ae_j=0 \quad (j=1, 2, \cdots, n).$$

由(5)式,Ae_j 是 A 的第 j 列,故 $A=O$。 证毕。

例 4 证明中的推理方法是常用的,须加以注意,并学会掌握它。

例 5 称 n 阶阵

$$F=\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}$$

为 **Frobenius 阵**。证明:对适合不等式 $1\leqslant s<n$ 的任一正整数 s 以及任一组不全为零的数 $b_0, b_1, \cdots, b_s$,必有

$$b_sF^s+b_{s-1}F^{s-1}+\cdots+b_1F+b_0I\neq O.$$

【证明】 记 $A=b_sF^s+b_{s-1}F^{s-1}+\cdots+b_1F+b_0I$,下面证明 A 中至少有一列非零。令

$$\alpha'=(-a_n, -a_{n-1}, \cdots, -a_1),$$

则

$$F=(e_2, e_3, \cdots, e_n, \alpha).$$

由(5)式,A 的第 1 列可表示为 Ae_1,而

$$Ae_1=b_sF^se_1+b_{s-1}F^{s-1}e_1+\cdots+b_1Fe_1+b_0e_1, \tag{9}$$

再利用(5)式可得

$$Fe_1=e_2, F^2e_1=Fe_2=e_3, \cdots, F^se_1=e_{s+1} \quad (s+1\leqslant n).$$

代入(9)式得到

$$\begin{aligned} Ae_1&=b_se_{s+1}+b_{s-1}e_s+\cdots+b_1e_2+b_0e_1 \\ &=(b_0, b_1, \cdots, b_s, 0, \cdots, 0)'. \end{aligned}$$

因 $b_0, b_1, \cdots, b_s$ 不全为零,故 Ae_1(即 A 的第 1 列)非零,于是

$A \neq O$。 证毕。

对照例 4、例 5 可见，欲证矩阵 $A=O$，可证 A 中任一列(行)是零；欲证 $A\neq O$，可证 A 中至少有一列(行)非零。

四、子矩阵

命题 1 说明，通过标准单位向量可将 A 与 A 的每一列（行）联系起来。进一步，下面将说明，通过标准单位向量可将 A 与它的任何子矩阵联系起来。

定义 在 $m\times n$ 阵 A 中，任取 k 行、l 列 ($1\leqslant k\leqslant m$, $1\leqslant l\leqslant n$)，位于这 k 行 l 列交叉点上的元素，按原来在 A 中的相对位置构成 $k\times l$ 阵，称此矩阵为 A 的一个**子矩阵**。

设 $A=(a_{ij})_{m\times n}$，由 A 中第 $i_1, i_2, \cdots, i_k$ 行、第 $j_1, j_2, \cdots, j_l$ 列 ($1\leqslant i_1<i_2<\cdots<i_k\leqslant m$；$1\leqslant j_1<j_2<\cdots<j_l\leqslant n$) 构成的子矩阵为

$$\begin{pmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_l} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_l} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_kj_1} & a_{i_kj_2} & \cdots & a_{i_kj_l} \end{pmatrix},$$

当 $k=l$ 时，称它为 k **阶子矩阵**，记为

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}=\begin{pmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_k} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_kj_1} & a_{i_kj_2} & \cdots & a_{i_kj_k} \end{pmatrix}$$

称 $A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}$ 为 k **阶主子阵**，称 $A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}$ 为 k **阶顺序主子阵**。

命题 2 设 $A=(a_{ij})_{m\times n}$，则

$$\begin{pmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_l} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_l} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_kj_1} & a_{i_kj_2} & \cdots & a_{i_kj_l} \end{pmatrix}=\begin{pmatrix} e'_{i_1} \\ e'_{i_2} \\ \vdots \\ e'_{i_k} \end{pmatrix} A(e_{j_1}, e_{j_2}, \cdots, e_{j_l})。 \quad (10)$$

【证明】 先证 $k=l=1$ 的情形，即

$$a_{ij}=e_i'Ae_j.$$

事实上

$$e_i'Ae_j=(a_{i1},a_{i2},\cdots,a_{ij},\cdots,a_{in})\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}=a_{ij}.$$

再利用(11)式证明(10)式：

$$\begin{pmatrix}e_{i_1}'\\ e_{i_2}'\\ \vdots\\ e_{i_k}'\end{pmatrix}A\,(e_{j_1},e_{j_2},\cdots,e_{j_l})$$

$$=\begin{pmatrix}e_{i_1}'Ae_{j_1} & e_{i_1}'Ae_{j_2} & \cdots & e_{i_1}'Ae_{j_l}\\ e_{i_2}'Ae_{j_1} & e_{i_2}'Ae_{j_2} & \cdots & e_{i_2}'Ae_{j_l}\\ \cdots & \cdots & \cdots & \cdots\\ e_{i_k}'Ae_{j_1} & e_{i_k}'Ae_{j_2} & \cdots & e_{i_k}'Ae_{j_l}\end{pmatrix}$$

$$=\begin{pmatrix}a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_l}\\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_l}\\ \cdots & \cdots & \cdots & \cdots\\ a_{i_kj_1} & a_{i_kj_2} & \cdots & a_{i_kj_l}\end{pmatrix}$$

证毕。

作为命题 2 的特款，应有

$$A\begin{pmatrix}i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}=\begin{pmatrix}e_{i_1}'\\ e_{i_2}'\\ \vdots\\ e_{i_k}'\end{pmatrix}A(e_{i_1},e_{i_2},\cdots,e_{i_k}).\tag{12}$$

$$A\begin{pmatrix}1 & 2 & \cdots & k\\ 1 & 2 & \cdots & k\end{pmatrix}=\begin{pmatrix}e_1'\\ e_2'\\ \vdots\\ e_k'\end{pmatrix}A(e_1, e_2, \cdots, e_k). \tag{13}$$

习　　题

1. 设A为正交阵(或对称阵、实镜象阵)，求证$M=\begin{pmatrix}I & O\\ O & A\end{pmatrix}$也是正交阵(或对称阵、实镜象阵)。

2. 设$A=[k_1I_{n_1}, k_2I_{n_2}, \cdots, k_sI_{n_s}]$，其中$\sum_{i=1}^{s} n_i=n$且$k_1, k_2, \cdots, k_s$互不相同。求证$n$阶方阵$B$与$A$可交换的充要条件是$B=[B_1, B_2, \cdots, B_s]$，其中$B_i$为$n_i$阶方阵$(i=1,2,\cdots,s)$。

3. 称复方阵A为 Hermite 阵(H－阵)，如果$A=\bar{A}'$。设复方阵A有分解式$A=B+\sqrt{-1}C$，其中B、C为实方阵。求证A为H－阵的充分必要条件是方阵$\begin{bmatrix}B & C\\ -C & B\end{bmatrix}$为对称阵。

4. 设U、V为n阶正交阵，D为r阶非异阵，且$r<n$。证明，如果

$$U\begin{pmatrix}D & O\\ O & O\end{pmatrix}U'=V\begin{pmatrix}D & O\\ O & O\end{pmatrix}V',$$则

$$U\begin{pmatrix}D^{-1} & O\\ O & O\end{pmatrix}U'=V\begin{pmatrix}D^{-1} & O\\ O & O\end{pmatrix}V'.$$

5. 设A、C为非异阵，求证$P=\begin{pmatrix}A & O\\ B & C\end{pmatrix}$为非异阵，并求出$P^{-1}$。

6. 证明n阶排列阵

$$J_n=\begin{pmatrix} & & 1\\ & ⋰ & \\ 1 & & \end{pmatrix}$$

是正交、对称阵。

7. 设

$$A=\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&0&1\\1&0&0&\cdots&0&0\end{pmatrix}_{n\times n}$$

求证

$$A^k=\begin{pmatrix}O&I_{n-k}\\I_k&O\end{pmatrix}\qquad(k=1,2,\cdots,n-1);A^n=I.$$

8. 设A、B均为$m\times n$阵. 证明:若对任一$n\times p$阵C, 恒有$AC=BC$, 则$A=B$。

9. 设A为$m\times n$阵, b为m维列向量。若对任何非零的n维列向量x, 恒有$Ax=b$,求证$A=O,b=0$。

10. 证明: 若矩阵A非零, 则必存在非零的列向量α和β, 使得$\alpha'A\beta\neq0$.

11. 证明:若n阶实方阵A, 对任一实的n维列向量x, 恒有$x'Ax=0$, 则A为反对称阵(即$A=-A'$)。

12. 以E_{ij}表示i行j列的元素为1、其余元素为零的方阵. 证明:

(i) $E_{ij}=e_ie_j'$, $E_{ij}E_{kl}=\delta_{jk}E_{il}$;

(ii) 若方阵A适合条件$AE_{ij}=E_{ij}A$, 则$a_{ii}=a_{jj}$, $a_{ki}=0\quad(k\neq i)$, $a_{jk}=0\quad(k\neq j)$。其中a_{ij}为A的(i,j)元素;

(iii) 与所有方阵可交换的方阵必定是纯量阵。

13. 写出下列矩阵的子矩阵与原矩阵的关系式:

(i)

$$A=\begin{pmatrix}1&2&-1&3&5\\1&-1&0&4&7\\0&5&1&5&6\\1&0&2&1&-1\end{pmatrix},$$

子矩阵$\begin{pmatrix}-1&4&7\\0&1&-1\end{pmatrix}$;

(ii) $A=(a_{ij})_{4\times 5}$,三阶主子阵$A\begin{pmatrix}1\ 3\ 4\\1\ 3\ 4\end{pmatrix}$.

14. 设A为$m\times n$阵。证明A与A'的子矩阵之间有如下关系:

$$\left(A\begin{pmatrix}i_1,i_2,\cdots,i_k\\j_1,j_2,\cdots,j_k\end{pmatrix}\right)'=A'\begin{pmatrix}j_1,j_2,\cdots,j_k\\i_1,i_2,\cdots,i_k\end{pmatrix}.$$

§5 初等变换与初等阵

矩阵的初等变换起源于解线性方程组的消元法(消去法)．然而，它的作用并不只限于解线性方程组．利用初等变换将矩阵 A 化为形状简单的矩阵 B，通过 B 来探讨 A 的某些性质，乃是矩阵论中常用的一种方法．这方面的大部分内容将安排在第三、八两章．本节仅介绍初等变换在化简矩阵以及求逆阵中的应用．

一、初等变换与初等阵

定义 分别称下列三种类型的变换为矩阵的第一、二、三种初等行(列)变换：

(1) 对调矩阵中任意两行(列)的位置；

(2) 以一非零常数乘矩阵中某一行(列)；

(3) 将矩阵中某一行(列)的 k 倍(k 为一个数)加到另一行(列)上去．

矩阵的初等行变换与初等列变换统称为矩阵的初等变换．

对调单位阵的第 i 行(列)与第 j 行(列)的位置，得到方阵

$$P_{ij}=\begin{array}{c}\begin{array}{ccccccccc}&&&i&&j&&&\end{array}\\\begin{pmatrix}1&&&&&&&\\&\ddots&&&&&&\\&&1&&&&&\\&&&0&\cdots&1&&\\&&&\vdots&\ddots&\vdots&&\\&&&1&\cdots&0&&\\&&&&&&1&\\&&&&&&&\ddots&\\&&&&&&&&1\end{pmatrix}\end{array}\begin{matrix}\\\\\\i\\\\j\\\\\\\\\end{matrix}, \tag{1}$$

以非零常数 c 乘单位阵的第 i 行(列)，得到方阵

$$D_i(c)=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&c&&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}i,\quad c\neq 0;\tag{2}$$

将单位阵中第 j 行(i 列)的 k 倍加到第 i 行(j 列),得到方阵

$$T_{ij}(k)=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&\cdots&k&&\\&&&\ddots&\vdots&&\\&&&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{matrix}\\ \\ i\\ \\ \\ \\ \\ \end{matrix}.\tag{3}$$

(其中 k 位于第 j 列。)

定义 分别称上述(1)、(2)、(3)种类型的方阵为第一、二、三种初等阵,并统称它们为初等阵.

命题 1 矩阵 A 作一次初等行(列)变换后所得的矩阵 B 等于以一个相应的初等阵左(右)乘 A.

【证明】 对于初等行变换,要证明的结论是:

(i) 若对调 A 中第 i 行与第 j 行的位置得到 B,记为

$$A=\begin{pmatrix}\vdots\\ \tilde{\alpha}_i\\ \vdots\\ \tilde{\alpha}_j\\ \vdots\end{pmatrix}\longrightarrow\begin{pmatrix}\vdots\\ \tilde{\alpha}_j\\ \vdots\\ \tilde{\alpha}_i\\ \vdots\end{pmatrix}=B,$$

则 $B=P_{ij}A$;

(ii) 若以数 $c(c\neq 0)$ 乘 A 的第 i 行得到 B,记为

$$A = c \to \begin{pmatrix} \vdots \\ \tilde{\alpha}_i \\ \vdots \end{pmatrix} \longrightarrow \begin{pmatrix} \vdots \\ c\tilde{\alpha}_i \\ \vdots \end{pmatrix} = B, \ c \neq 0,$$

则 $B = D_i(c)A$；

(iii) 若将 A 中第 j 行的 k 倍加到第 i 行上去得到 B，记为

$$A = \begin{matrix} \to \\ \\ k \end{matrix} \begin{pmatrix} \vdots \\ \tilde{\alpha}_i \\ \vdots \\ \tilde{\alpha}_j \\ \vdots \end{pmatrix} \longrightarrow \begin{pmatrix} \vdots \\ \tilde{\alpha}_i + k\tilde{\alpha}_j \\ \vdots \\ \tilde{\alpha}_j \\ \vdots \end{pmatrix} = B,$$

则 $B = T_{ij}(k)A$，其中，$\tilde{\alpha}_i$ 代表 A 的第 i 行。

设 A 为 $m \times n$ 阵，则左乘的初等阵的阶数应是 m。注意到

$$P_{ij} = \begin{pmatrix} e_1' \\ \vdots \\ e_j' \\ \vdots \\ e_i' \\ \vdots \\ e_m' \end{pmatrix}, D_i(c) = \begin{pmatrix} e_1' \\ \vdots \\ ce_i' \\ \vdots \\ e_m' \end{pmatrix}, \ c \neq 0;$$

$$T_{ij}(k) = \begin{pmatrix} e_1' \\ \vdots \\ e_i' + ke_j' \\ \vdots \\ e_j' \\ \vdots \\ e_m' \end{pmatrix},$$

并利用分块乘法即可证明(i)—(iii)。以(iii)为例，即是

$$
T_{ij}(k)A=\begin{pmatrix} e_1'A \\ \vdots \\ (e_i'+ke_j')A \\ \vdots \\ e_j'A \\ \vdots \\ e_m'A \end{pmatrix}=\begin{pmatrix} \tilde{\alpha}_1 \\ \vdots \\ \tilde{\alpha}_i+k\tilde{\alpha}_j \\ \vdots \\ \tilde{\alpha}_j \\ \vdots \\ \tilde{\alpha}_m \end{pmatrix}=B.
$$

若将初等阵按列分块，则对初等列变换，类似地可证明下列结论：

(i)′ 若 $A=(\cdots\alpha_i\cdots\alpha_j\cdots)\longrightarrow(\cdots\overset{i}{\alpha_j}\cdots\overset{j}{\alpha_i}\cdots)=B$，则 $B=AP_{ij}$；

(ii)′ 若 $A=(\cdots\overset{c}{\alpha_i}\cdots)\rightarrow(\cdots\overset{i}{c\alpha_i}\cdots)=B$，$c\neq0$，则 $B=AD_i(c)$；

(iii)′ 若 $A=(\cdots\alpha_i\cdots\alpha_j\cdots)\rightarrow(\cdots\alpha_i\cdots\overset{j}{k\alpha_i+\alpha_j}\cdots)=B$，则 $B=AT_{ij}(k)$，其中，α_j 代表 A 的第 j 列。 证毕。

为了便于记忆，我们将命题 1 所述的初等阵乘矩阵与矩阵作初等变换之间的关系概括为：**“左行右列、首尾为主”的八字规则**。“左行右列”意即，**左**乘初等阵的结果相当于在矩阵中作相应的**行**变换；**右**乘初等阵的结果相当于在矩阵中作相应的**列**变换。“首尾为主”是专对第三种初等阵来说的，即 $T_{ij}(k)$ 左乘 A 致使 A 的第 i 行变化，“i”位于 $T_{ij}(k)$ 的下标之**首**；$T_{ij}(k)$ 右乘 A 致使 A 的第 j 列变化，“j”位于 $T_{ij}(k)$ 的下标之**尾**。

由命题 1 不难证明，$P_{ij}P_{ij}=I$；$D_i(c)D_i(c^{-1})=I$，$c\neq0$；$T_{ij}(k)T_{ij}(-k)=T_{ij}(-k)T_{ij}(k)=I$。

命题 2 初等阵都是非异阵，且其逆阵为同类型的初等阵。即 $P_{ij}^{-1}=P_{ij}$；$D_i^{-1}(c)=D_i(c^{-1})$，$c\neq0$；$T_{ij}^{-1}(k)=T_{ij}(-k)$。

作为练习，读者容易证明，非异阵经初等变换后仍为非异阵；**奇异阵经初等变换后仍为奇异阵**。

我们约定，初等行（列）变换的示意箭头标在原矩阵的左（上）边，如命题 1 证明中所示。

例1 将

$$A=\begin{pmatrix} 1 & 2 & 3 & 1 \\ -1 & -1 & 2 & 3 \\ 1 & 0 & -1 & 4 \end{pmatrix}$$

的第1行减去第3行得到 B，即

$$A=\begin{matrix} \rceil\!\!\rightarrow \\ | \\ (-1) \end{matrix}\begin{pmatrix} 1 & 2 & 3 & 1 \\ -1 & -1 & 2 & 3 \\ 1 & 0 & -1 & 4 \end{pmatrix}$$

$$\rightarrow\begin{pmatrix} 0 & 2 & 4 & -3 \\ -1 & -1 & 2 & 3 \\ 1 & 0 & -1 & 4 \end{pmatrix}=B,$$

由八字规则可得：$B=T_{13}(-1)A$。亦即

$$\begin{pmatrix} 0 & 2 & 4 & -3 \\ -1 & -1 & 2 & 3 \\ 1 & 0 & -1 & 4 \end{pmatrix}$$

$$=\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 1 \\ -1 & -1 & 2 & 3 \\ 1 & 0 & -1 & 4 \end{pmatrix}.$$

二、用初等变换化矩阵为对角形

定理5.1 任一 $m\times n$ 阵 A，必可经有限次初等变换化为如下形式的 $m\times n$ 对角形矩阵①

---

① 矩阵

$$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{m\times n} \qquad (m>n)$$

以及

$$\begin{pmatrix} a_1 & & & & 0 & \vdots & 0 \\ & a_2 & & & 0 & \vdots & 0 \\ & & \ddots & & \vdots & \vdots & \vdots \\ & & & a_m & 0 & \vdots & 0 \end{pmatrix}_{m\times n} \qquad (m<n)$$

称为对角形矩阵.

$$\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}.$$

亦即存在初等阵 $P_1, P_2, \cdots, P_s$ 及 $Q_1, Q_2, \cdots, Q_t$，使得

$$P_s\cdots P_2P_1AQ_1Q_2\cdots Q_t = \begin{pmatrix} I_s & O \\ O & O \end{pmatrix},$$

其中，s 是适合 $0\leqslant s\leqslant \min(m, n)$ 的整数①；并约定 I_0 代表零矩阵.

【证明】 设 $A=(a_{ij})_{m\times n}$.

若 $A=0$，则 A 已是结论中所述的形式；

若 $A\neq 0$，总可用第一种初等变换将 A 变为左上角元素非零的矩阵，故可设 $a_{11}\neq 0$. 将第 1 行的 $(-a_{11}^{-1}a_{i1})$ 倍加到第 i 行上去 $(i=2, 3, \cdots, m)$，将第 1 列的 $(-a_{11}^{-1}a_{1j})$ 倍加到第 j 列上去 $(j=2, 3, \cdots, n)$，并以 a_{11}^{-1} 乘第 1 行，可将 A 变为如下形式

$$\begin{pmatrix} 1 & 0 \\ 0 & A_1 \end{pmatrix},$$

其中，A_1 为 $(m-1)\times(n-1)$ 阵.

对 A_1 重复以上的讨论………，直至所得的矩阵是零阵，便可得到定理的结论. 证毕.

本定理中的 s 由矩阵 A 唯一确定(见本章选做题 13). s 的重要含义将在第三章中说明.

三、非异阵与初等阵的关系、逆阵的求法

定理 5.2 A 为非异阵的充分必要条件是 A 可表示为有限个初等阵的乘积.

【证明】 充分性是显然的. 下面证明必要性.

设 A 的阶数为 n，由定理 5.1，存在初等阵 $P_1, P_2, \cdots, P_s$ 以及 $Q_1, Q_2, \cdots, Q_t$，使得

$$P_s\cdots P_2P_1AQ_1Q_2\cdots Q_t = \begin{pmatrix} I_s & O \\ O & O \end{pmatrix}, s\leqslant n.$$

① 记号 $\min(m, n)$ 表示 m、n 中较小的那个数.

因 A 非异，故 $\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}$ 也非异，从而必定有 $s=n$. 否则，$\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}$ 中至少有一行全为零而成为奇异阵(§3 习题 1)。于是得到

$$P_s\cdots P_2P_1AQ_1Q_2\cdots Q_t=I_n,$$

亦即

$$A=P_1^{-1}P_2^{-1}\cdots P_s^{-1}Q_t^{-1}\cdots Q_2^{-1}Q_1^{-1}.$$

显然，P_i^{-1}、Q_j^{-1} 是初等阵($i=1, 2, \cdots, s; j=1, 2, \cdots, t$)。

证毕。

利用定理 5.2 可将定理 5.1 写成下面的等价形式：

定理 5.1′ 对任一 $m\times n$ 阵 A，必存在 m 阶非异阵 P、n 阶非异阵 Q，使得

$$PAQ=\begin{pmatrix} I_s & O \\ O & O \end{pmatrix},$$

其中，s 是适合 $0\leqslant s\leqslant \min(m, n)$的整数。

此处"等价"的含义是，由定理 5.1 可推得定理 5.1′，反之亦然。由定理 5.1′，容易得到下面的推论。

推论 任一 n 阶方阵 A，必存在分解式

$$A=L\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}M, \tag{4}$$

其中，L、M 为非异阵，$s\leqslant n$。

命题 3 设 A、B 为 n 阶方阵。若 $AB=I_n$（或 $BA=I_n$），则 A、B 均为非异阵，且它们互为逆阵。

【证明】 由上述推论可知，存在非异阵 L、M，使得

$$A=L\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}M,\ s\leqslant n.$$

因 $AB=I_n$，所以

$$L\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}MB=I_n.$$

上式左乘 L^{-1} 得到

$$\begin{pmatrix} I_s & O \\ O & O \end{pmatrix} MB = L^{-1},$$

由此即可断定 $s=n$（否则，上式左边矩阵至少有一行是零而成为奇异阵，此与 L^{-1} 非异矛盾），从而 $A=LM$。据§3 命题 3，可知 A 为非异阵。

在等式 $AB=I_n$ 上左乘 A^{-1} 可得 $B=A^{-1}$，故 B 亦为非异阵，且易知 $B^{-1}=A$。 证毕。

命题 3 表明，$AB=I$ 的充分必要条件是 $BA=I$。故今后一旦验证了 $AB=I$（或 $BA=I$），便能断定 A 与 B 均非异。

例 2 设 A、C 分别为 m、n 阶方阵，B 为 $m\times n$ 阵，又 $M=\begin{pmatrix} A & B \\ O & C \end{pmatrix}$ 为非异阵，求证，A、C 均为非异阵。

【证明】 注意到 M 为非异阵，将 M^{-1} 作与 M 相同的分块 $M^{-1}=\begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$，则

$$MM^{-1}=\begin{pmatrix} * & * \\ * & CW \end{pmatrix}=\begin{pmatrix} I_m & O \\ O & I_n \end{pmatrix},$$

于是得到

$$CW=I_n.$$

由命题 3 可知，C 为非异阵。再利用 $M^{-1}M=I_{m+n}$，同理可证 A 亦为非异阵。 证毕。

命题 4 方阵 A 非异的充分必要条件是，对 A 作有限次初等行（列）变换后可将 A 化为单位阵。

【证明】 若 A 非异，则 $A^{-1}A=I$。因 A 非异时 A^{-1} 也非异，据定理 5.2 可设

$$A^{-1}=P_s\cdots P_2P_1,$$

其中，$P_i(i=1,2,\cdots,s)$ 为初等阵。于是，$A^{-1}A=I$，即为

$$P_s\cdots P_2P_1A=I.$$

上式表明，对 A 作有限次初等行变换后可将 A 化为单位阵，从而必要性得证。

反之，若对 A 作有限次初等行变换后可将 A 化为单位阵，即存在初等阵 $P_i(i=1, 2, \cdots, s)$，使得

$$P_s\cdots P_2P_1A=I,$$

则由命题 3 可知，A 为非异阵，且 $A^{-1}=P_s\cdots P_2P_1$。 证毕。

命题 4 的充分性，给出了利用初等变换求逆阵的方法：若对分块阵 (A, I) 只作初等行变换后能化成 (I, B) 的形式，则 A 为非异阵，且 $A^{-1}=B$。

例 3 证明方阵

$$A=\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ -1 & 4 & 3 \end{pmatrix}$$

是非异阵，并求其逆阵。

【解】

$$(A, I_3)=\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 1 & 0 \\ -1 & 4 & 3 & 0 & 0 & 1 \end{array}\right)\to$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right)\to$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & -3 & -4 & 1 \end{array}\right)\to$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & -3 & 1 \\ 0 & 0 & -1 & -3 & -4 & 1 \end{array}\right)$$

$$\to\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & -3 & 1 \\ 0 & 0 & 1 & 3 & 4 & -1 \end{array}\right),$$

（变换说明：第一步将第 1 行分别加到第 2、3 行；第二步将第 2 行乘 (-4) 加到第 3 行；第三步将第 3 行乘 1 加到第 2 行；第四步将第 3 行乘 (-1)。）

由命题 4 可知，A 非异，且

$$A^{-1}=\begin{pmatrix}1 & 0 & 0\\ -2 & -3 & 1\\ 3 & 4 & -1\end{pmatrix}.$$

现在,回过来讨论一下定理5.1′中非异阵P及Q的求法. 对于给定的A,在用初等变换化为$\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}$时,只要记下相应于初等行变换的初等阵$P_1, P_2, \cdots, P_s$以及相应于初等列变换的初等阵$Q_1, Q_2, \cdots, Q_t$,令$P=P_s\cdots P_2P_1$, $Q=Q_1Q_2\cdots Q_t$,即可求得使$PAQ=\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}$成立的非异阵$P$以及$Q$. 但是,这一方法相当麻烦. 下面给出利用初等变换同时求得P、Q的方法. 这种方法无需记下相应的初等阵,也不必求它们的乘积,其基础是下面的命题.

命题5 设A、P、Q为定理5.1′中所述,则

$$\begin{pmatrix}P & O\\ O & I_n\end{pmatrix}\begin{pmatrix}A & I_m\\ I_n & O\end{pmatrix}\begin{pmatrix}Q & O\\ O & I_m\end{pmatrix}=\begin{pmatrix}\begin{pmatrix}I_r & O\\ O & O\end{pmatrix} & P\\ Q & O\end{pmatrix}. \tag{5}$$

【证明】 作分块乘法,便可验证(5)式. 证毕.

因P、Q非异,故$\begin{pmatrix}P & O\\ O & I_n\end{pmatrix}$及$\begin{pmatrix}Q & O\\ O & I_m\end{pmatrix}$也非异. 由定理5.2,可将它们表示为有限个初等阵的乘积. 于是,(5)式表明,只要对分块矩阵

$$\begin{pmatrix}A & I_m\\ I_n & O\end{pmatrix}$$

的前m行、前n列施行有限次初等变换化为

$$\begin{pmatrix}\begin{pmatrix}I_r & O\\ O & O\end{pmatrix} & P\\ Q & O\end{pmatrix}$$

的形式，则 P、Q 就是使得 $PAQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 成立的一对非异阵。

例 4 设

$$A=\begin{pmatrix} 0 & 1 & -1 & -1 \\ 1 & 1 & 0 & 1 \\ -1 & -2 & 1 & 0 \end{pmatrix}.$$

求非异阵 P 及 Q，使得 PAQ 成为 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的形式。

【解】

$$\begin{pmatrix} A & I_3 \\ I_4 & 0 \end{pmatrix}=\left(\begin{array}{cccc|ccc} 0 & 1 & -1 & -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ -1 & -2 & 1 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & & 0 & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \end{array}\right) \to$$

$$\xrightarrow{1}\left(\begin{array}{cccc|ccc} 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 & 0 \\ -1 & -2 & 1 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & & 0 & \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \end{array}\right)$$

$$
\left(\begin{array}{cccc|ccc}
1 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & -1 & -1 & 1 & 0 & 0 \\
0 & -1 & 1 & 1 & 0 & 1 & 1 \\
\hline
1 & 0 & 0 & 0 & & & \\
0 & 1 & 0 & 0 & & 0 & \\
0 & 0 & 1 & 0 & & & \\
0 & 0 & 0 & 1 & & &
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|ccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & -1 & -1 & 1 & 0 & 0 \\
0 & -1 & 1 & 1 & 0 & 1 & 1 \\
\hline
1 & -1 & 0 & -1 & & & \\
0 & 1 & 0 & 0 & & 0 & \\
0 & 0 & 1 & 0 & & & \\
0 & 0 & 0 & 1 & & &
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|ccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & -1 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 \\
\hline
1 & -1 & 0 & -1 & & & \\
0 & 1 & 0 & 0 & & 0 & \\
0 & 0 & 1 & 0 & & & \\
0 & 0 & 0 & 1 & & &
\end{array}\right)
$$

$$
\to \left(\begin{array}{cccc:ccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 \\
\hdashline
1 & -1 & -1 & -2 & & & \\
0 & 1 & 1 & 1 & & O & \\
0 & 0 & 1 & 0 & & & \\
0 & 0 & 0 & 1 & & &
\end{array}\right)
$$

得到

$$
P=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad Q=\begin{pmatrix} 1 & -1 & -1 & -2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
$$

四、分块初等阵及其应用

定义 分别称下列三种类型的变换为分块矩阵的第一、二、三种初等行(列)变换：

(1) 对调分块矩阵中任意两行(列)的位置；

(2) 以一非异阵左(右)乘分块矩阵中某一行(列)；

(3) 以矩阵 K 左(右)乘分块矩阵中某一行(列)后加到另一行(列)上去。

分块矩阵的初等行变换与初等列变换统称为分块矩阵的初等变换。

对分块的 t 阶单位阵 $I=[I_{t_1}, I_{t_2}, \cdots, I_{t_s}]$，对调其第 i 行与第 j 行的位置，得到分块阵

$$
P_{ij}(I_{t_i}, I_{t_j}) = \begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}
\begin{pmatrix}
I_{t_1} & & & & \\
& \ddots & & & \\
& & O & \cdots & I_{t_j} & \\
& & \vdots & \ddots & \vdots & \\
& & I_{t_i} & \cdots & O & \\
& & & & & \ddots \\
& & & & & & I_{t_s}
\end{pmatrix} \quad (i<j),
$$

对调其第 i 列与第 j 列的位置，得到分块阵

$$P_{ij}(I_{l_i}, I_{l_j}) = \begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\begin{pmatrix} I_{l_1} & & & & \\ & \ddots & & & \\ & & 0 & \cdots & I_{l_i} & & \\ & & \vdots & \ddots & \vdots & & \\ & & I_{l_j} & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & I_{l_t} \end{pmatrix} \quad (i<j);$$

以非异阵 C 左(右)乘其第 i 行(列)，得到分块阵

$$D_i(C) = [I_{l_1}, \cdots, I_{l_{i-1}}, C, I_{l_{i+1}}, \cdots, I_{l_t}];$$

以矩阵 K 左(右)乘其第 j 行(i 列)后加到第 i 行(j 列)上去，得到分块矩阵

$$T_{ij}(K) = \begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\begin{pmatrix} I_{l_1} & & & & \\ & \ddots & & & \\ & & I_{l_i} & \cdots & K & \\ & & & \ddots & \vdots & \\ & & & & I_{l_j} & \\ & & & & & \ddots & \\ & & & & & & I_{l_t} \end{pmatrix}.$$

定义 分别称上述 $P_{ij}(I_{l_j}, I_{l_i})$(或 $P_{ij}(I_{l_i}, I_{l_j})$)、$D_i(C)$、$T_{ij}(K)$这三种类型的分块阵为第一、二、三种**分块初等阵**，并统称它们为**分块初等阵**。

设 $A=(A_{ij})_{r\times s}$，其中 A_{ij} 为 $m_i\times n_j$ 阵 ($i=1, 2, \cdots, r; j=1, 2, \cdots, s$)，则当 $P_{ij}(I_{l_j}, I_{l_i})$、$D_i(C)$、$T_{ij}(K)$ 左乘A时，须取 $t=r$, $l_i=m_i(i=1, 2, \cdots, r)$；当 $P_{ij}(I_{l_i}, I_{l_j})$、$D_i(C)$、$T_{ij}(K)$ 右乘A时，须取 $t=s$, $l_j=n_j(j=1, 2, \cdots, s)$。容易证明，对分块矩阵$A$作一次初等行(列)变换后所得的分块阵$B$等于以一个相应的分块初等阵左(右)乘 A，且“八字规则”仍旧适用(类似命题 1. 但需注意，在对调行时，相应的分块初等阵取 $P_{ij}(I_{l_j}, I_{l_i})$；在对调列时，相应的分块初等阵取 $P_{ij}(I_{l_i}, I_{l_j})$)。

根据命题4，若对分块阵A作有限次初等行(列)变换后得到单位阵，则A是非异阵，且A^{-1}是相应于这些初等变换的分块初等阵的积. 这一结果，亦可用于求分块阵的逆阵，其演算形式类似于例 3。

例 5 设 A、C 分别是 m、n 阶非异阵，B 是 $n\times m$ 阵，求证 $M=\begin{pmatrix} O & A \\ C & B \end{pmatrix}$ 是非异阵，并求出 M^{-1}。

【解】 将 (M, I_{n+m}) 用分块阵的初等行变换化为 (I_{n+m}, R)，则 $M^{-1}=R$。

$$(M, I_{n+m})=\begin{pmatrix} O & A & \vdots & I_m & O \\ C & B & \vdots & O & I_n \end{pmatrix}\to\begin{pmatrix} C & B & \vdots & O & I_n \\ O & A & \vdots & I_m & O \end{pmatrix}$$

$$\xrightarrow{(-BA^{-1})}\begin{pmatrix} C & O & \vdots & -BA^{-1} & I_n \\ O & A & \vdots & I_m & O \end{pmatrix}\xrightarrow{C^{-1},\ A^{-1}}\begin{pmatrix} I_n & O & \vdots & -C^{-1}BA^{-1} & C^{-1} \\ O & I_m & \vdots & A^{-1} & O \end{pmatrix}.$$

所以

$$\begin{pmatrix} O & A \\ C & B \end{pmatrix}^{-1}=\begin{pmatrix} -C^{-1}BA^{-1} & C^{-1} \\ A^{-1} & C \end{pmatrix}.$$

特别，当 $B=O$ 时，得到

$$\begin{pmatrix} O & A \\ C & O \end{pmatrix}^{-1}=\begin{pmatrix} O & C^{-1} \\ A^{-1} & O \end{pmatrix}.$$

例 6 (**Sherman–Morrison** 公式)设 A 是 n 阶非异阵，α、β 是 n 维列向量，且 $\beta' A^{-1}\alpha\neq -1$，则 $A+\alpha\beta'$ 是非异阵，且

$$(A+\alpha\beta')^{-1}=A^{-1}-\frac{A^{-1}\alpha\beta' A^{-1}}{1+\beta' A^{-1}\alpha}.$$

【证明】 下面给出这一公式的由来。它是将 $A+\alpha\beta'$ "升阶"后再利用分块阵借初等变换求逆阵的方法而推导出来的。

$$\begin{pmatrix} 1 & -\beta' & \vdots & 1 & O \\ O & A+\alpha\beta' & \vdots & O & I_n \end{pmatrix}\xrightarrow{(\alpha)}\ \xrightarrow{(\beta' A)}\begin{pmatrix} 1 & -\beta' & \vdots & 1 & O \\ \alpha & A & \vdots & \alpha & I_n \end{pmatrix}$$

$$\rightarrow \frac{1}{(1+\beta' A^{-1}\alpha)} \begin{pmatrix} 1+\beta' A^{-1}\alpha & O & \vdots & 1+\beta' A^{-1}\alpha & \beta' A^{-1} \\ \alpha & A & \vdots & \alpha & I_n \end{pmatrix}$$

$$\rightarrow \begin{matrix} (1+\beta' A^{-1}\alpha)^{-1}\rightarrow \\ A^{-1}\rightarrow \end{matrix} \begin{pmatrix} 1+\beta' A^{-1}\alpha & O & \vdots & 1+\beta' A^{-1}\alpha & \beta' A^{-1} \\ O & A & \vdots & O & I_n - \frac{\alpha\beta' A^{-1}}{1+\beta' A^{-1}\alpha} \end{pmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & O & \vdots & 1 & \frac{\beta' A^{-1}}{1+\beta' A^{-1}\alpha} \\ O & I_n & \vdots & O & A^{-1} - \frac{A^{-1}\alpha\beta' A^{-1}}{1+\beta' A^{-1}\alpha} \end{bmatrix}.$$

这说明$\begin{pmatrix} 1 & -\beta \\ O & A+\alpha\beta' \end{pmatrix}$是非异阵，因而 $A+\alpha\beta'$ 也是非异阵（由例 2），且

$$\begin{pmatrix} 1 & -\beta \\ O & A+\alpha\beta' \end{pmatrix}^{-1} = \begin{bmatrix} 1 & \frac{\beta' A^{-1}}{1+\beta' A^{-1}\alpha} \\ O & A^{-1} - \frac{A^{-1}\alpha\beta' A^{-1}}{1+\beta' A^{-1}\alpha} \end{bmatrix}.$$

但由§4 的例 1，应有

$$\begin{pmatrix} 1 & -\beta \\ O & A+\alpha\beta' \end{pmatrix}^{-1} = \begin{pmatrix} 1 & * \\ O & (A+\alpha\beta')^{-1} \end{pmatrix},$$

比较上面两式，即可得到所要证明的公式。　　　　　　证毕.

Sherman-Morrison 公式在计算数学及最优化理论中都有应用。

习　题

1．对下列初等变换，写出相应的初等阵以及 B 与 A 之间的关系式：

(i) $A = \begin{matrix} (-2) \\ \rightarrow \end{matrix} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 2 & 3 & 1 & 2 \\ 1 & -1 & 2 & 1 \end{pmatrix} \rightarrow \begin{matrix} 1\text{—— —}\downarrow \\ \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 3 & 3 & -2 \\ 1 & -1 & 2 & 1 \end{pmatrix} \end{matrix}$

$$\rightarrow\begin{pmatrix}1 & 0 & 0 & 2\\ 0 & 3 & 3 & -2\\ 1 & -1 & 3 & 1\end{pmatrix}=B;$$

(ii)

$$A=\begin{pmatrix}1 & -1 & 1\\ -1 & 1 & -1\\ 1 & -1 & 1\end{pmatrix}\rightarrow\begin{pmatrix}1 & 0 & 0\\ -1 & 0 & 0\\ 1 & 0 & 0\end{pmatrix}\rightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}=B.$$

(第一步：第1列乘 1 加到第2列，第1列乘 (-1) 加到第3列；第二步：第1行乘 1 加到第2行，第1行乘 (-1) 加到第3行。)

2．设A为非异阵．求证

(i) 若对调A的第i、j行位置得到B，则B^{-1}恰等于对调A^{-1}的第i、j列位置所得的矩阵；

(ii) 若以非零常数c乘A的第j列得到矩阵P，则P^{-1}恰等于A^{-1}的第j行除以c所得的矩阵．

3．证明：与所有的非异阵可交换的方阵必定是纯量阵．

4．设A为n阶非零奇异阵．证明，存在n阶非零奇异阵B、C，使得$AB=CA=O$．

5．证明：两个奇异阵之积仍为奇异阵．

6．分解下列方阵为初等阵的乘积：

(i) $A=\begin{pmatrix}1 & 0 & 0\\ -2 & 4 & 1\\ -1 & 3 & 1\end{pmatrix}$；　(ii) $A=\begin{pmatrix}2 & 3 & 4\\ 1 & -2 & 3\\ 3 & 1 & 8\end{pmatrix}$．

7．用初等变换求下列方阵的逆阵：

(i) $A=\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix}$；　(ii) $A=\begin{pmatrix}2 & 1 & 0 & 0\\ 0 & 2 & 1 & 0\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 2\end{pmatrix}$；

(iii) $A=\begin{pmatrix}0 & 0 & 1 & -1\\ 0 & 3 & 1 & 4\\ 2 & 7 & 6 & -1\\ 1 & 2 & 2 & -1\end{pmatrix}$；　(iv) $A=\begin{pmatrix}1 & 1 & \cdots & 1\\ & 1 & \cdots & 1\\ & & \ddots & \vdots\\ O & & & 1\end{pmatrix}_{n\times n}$．

8．求非异阵P及Q，使得 $PAQ=\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}$．

(i)
$$A=\begin{pmatrix}1 & -1\\ -1 & 0\\ 2 & 3\end{pmatrix},\qquad \text{(ii)}\quad A=\begin{pmatrix}3 & 1 & 2\\ -1 & 0 & 1\\ 2 & 1 & 0\end{pmatrix}.$$

9. 证明：第一种初等阵可表示为第二种与第三种初等阵的乘积。

10. 设 u、v 是 n 维列向量，记 $H(u, v)=I_n+uv'$。求证第一、二、三种初等阵均可写成形如 $H(u, v)$ 的矩阵。

11. 设

$$A=\begin{pmatrix}O & \frac{1}{2}B\\ \frac{1}{2}B & O\end{pmatrix},$$

其中 B 是方阵。仿照例 4 的方法，求出非异阵 P，使得 $P'AP$ 为分块对角阵。

12. 求下列方阵的逆阵：

(i)
$$A=\begin{pmatrix} & & & a_1\\ & & a_2 & \\ & \iddots & & \\ a_n & & & \end{pmatrix},\quad 其中\ \prod_{i=1}^{n}a_i\neq 0^{①},$$

(ii)
$$A=\begin{pmatrix}0 & a_1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & a_{n-1}\\ a_n & 0 & \cdots & 0\end{pmatrix},\quad 其中\ \prod_{i=1}^{n}a_i\neq 0;$$

(iii)
$$F=\begin{pmatrix}0 & 0 & \cdots & 0 & -a_n\\ 1 & 0 & \cdots & 0 & -a_{n-1}\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & -a_1\end{pmatrix},其中\ a_n\neq 0.$$

13. 设 A、B 为同阶实方阵，且 A 及 $C=A-\sqrt{-1}B$ 都是非异阵（显然，$\bar{C}=A+\sqrt{-1}B$ 也是非异阵）。求证：

---

① $\prod_{i=1}^{n}a_i=a_1a_2\cdots a_n$.

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix}^{-1}$$

$$=\begin{pmatrix} C^{-1}-\sqrt{-1}\overline{C}^{-1}BC^{-1} & \sqrt{-1}(C^{-1}-\overline{C}^{-1}-\sqrt{-1}\overline{C}^{-1}BC^{-1}) \\ \overline{C}^{-1}BC^{-1} & \overline{C}^{-1}+\sqrt{-1}\overline{C}^{-1}BC^{-1} \end{pmatrix}.$$

练 做 题

1. 设A为 $m\times n$ 复矩阵. 证明：若 $A\overline{A}'=O$(或 $\overline{A}'A=O$)，则$A=O$.

2. 证明：对任何 $1\leqslant k\leqslant n$，恒有

$$A_1+A_2+\cdots+A_n=(A_1+A_2+\cdots+A_k)+(A_{k+1}+\cdots+A_n),$$

其中，$A_i(i=1, 2, \cdots, n)$为同阶方阵. 称上式为加法的一般结合律.

3. 设A、B为同阶方阵，且 $AB=BA$. 求证：

$$(A+B)^n=A^n+C_n^1A^{n-1}B+C_n^2A^{n-2}B^2+\cdots+B^n=\sum_{k=0}^{n}C_n^kA^{n-k}B^k.$$

4. 证明：对任一实方阵A，必存在唯一的分解式 $A=B+C$，其中 $B=B'$，$C=-C'$.

5. 设A为 $m\times n$ 实矩阵，s 为 n 阶反对称阵. 证明：对任一实的 n 维列向量 x，成立

(i) $x'Ax=0$；

(ii) $x'(I_n+s)x>0$，且等号成立$\Longleftrightarrow x=0$；

(iii) $x'(A'A+S)x>0$. 又若等号成立，问 x 是否为零向量？

6. 设

$$A=\begin{pmatrix} a_1 & a_2 \\ -\overline{a_2} & \overline{a_1} \end{pmatrix},\quad B=\begin{pmatrix} b_1 & b_2 \\ -\overline{b_2} & \overline{b_1} \end{pmatrix},$$

证明：(i) AB 与 BA 也是形如

$$\begin{pmatrix} c_1 & c_2 \\ -\overline{c_2} & \overline{c_1} \end{pmatrix}$$

的方阵；

(ii) 若 a_i、$b_i(i=1, 2)$为实数，则 $AB=BA$；

(iii) 若 a_i、$b_i(i=1, 2)$为复数，则未必有 $AB=BA$.

7. 设A为非异阵，规定

$$A^{-k}=(A^{-1})^k \qquad (k \text{ 为正整数}).$$

证明：(i) $A^kA^l=A^{k+l}$；

(ii) $(A^k)^l=A^{kl}$；

(iii) 若 $AB=BA$，则 $(AB)^k=A^kB^k$；

其中，k、l 为整数（当方阵的负次幂出现时，认为该方阵是非异阵）。

8. 分别对适合下列等式的方阵 A，证明 $I-A$ 是非异阵，并求出 $(I-A)^{-1}$。

(i) $A^2=2A$；

(ii) $A^2-A+I=O$；

(iii) $A^3=3A(A-I)$；

(iv) $A^k=O$（k 为一正整数）。

9. 设 A 为 $n\times m$ 阵，B 为 $m\times n$ 阵。若 I_n-AB 是非异阵，证明 I_m-BA 也是非异阵，并求出 $(I_m-BA)^{-1}$。

（提示：应用关系式 $(I_n-AB)A=A(I_m-BA)$ 以及关系式 $B(I_n-AB)=(I_m-BA)B$，并注意“I_n-AB 是非异阵”的条件。）

10. 设 n 阶 Frobenius 矩阵

$$F=\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix}.$$

(i) 求证 $F^n+a_1F^{n-1}+\cdots+a_nI_n=O$；

（提示：记上式左边为 B，且注意 $BF=FB$）

(ii) 若 $A=(a_{ij})_{n\times n}$ 与 F 可交换，求证：

$$A=a_{n1}F^{n-1}+a_{n-1,1}F^{n-2}+\cdots+a_{21}F+a_{11}I_n.$$

11. 称方阵

$$\begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-3} & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \cdots & c_{n-4} & c_{n-3} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots & \cdots\cdots \\ c_2 & c_3 & c_4 & \cdots & c_0 & c_1 \\ c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_0 \end{pmatrix}$$

为 n 阶循环阵。证明：循环阵的乘积仍为循环阵。

（提示：利用 §4 习题 7 的结论。）

12. 方阵 $A=(a_{ij})_{n\times n}$ 称为上（下）三角阵，如果 $i>j\,(i<j)$ 时有 $a_{ij}=0$。证明：

(i) 两个上(下)三角阵的乘积仍为上(下)三角阵；

(ii) 非异上(下)三角阵的逆阵仍为上(下)三角阵。

13. 证明定理5.1中的 s 由 A 唯一确定。

(提示：若 s 不唯一，则必存在 m 阶非异阵 P、P_1 以及 n 阶非异阵 Q、Q_1，使得 $P\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}Q = A = P_1\begin{pmatrix} I_t & O \\ O & O \end{pmatrix}Q_1$，其中 $s \neq t$。上式可改写为

$(P_1^{-1}P)\begin{pmatrix} I_s & O \\ O & O \end{pmatrix} = \begin{pmatrix} I_t & O \\ O & O \end{pmatrix}(Q_1Q^{-1})$。对 Q_1Q^{-1} 适当分块并作乘法，利用§5例2的结论可知，"$s<t$"不可能.)

14. 设 A 是 m 阶非异阵，D 是 n 阶阵，B 与 C 分别是 $m\times n$ 与 $n\times m$ 阵。求证：$M\equiv\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 是非异阵 $\Longleftrightarrow D-CA^{-1}B$ 是非异阵；且当 $F\equiv D-CA^{-1}B$ 是非异阵时，成立

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1}+A^{-1}BF^{-1}CA^{-1} & -A^{-1}BF^{-1} \\ -F^{-1}CA^{-1} & F^{-1} \end{pmatrix}.$$

上式称为非异阵求逆的第一降阶定理。

$\left(\text{提示，用分块阵的初等变换将} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{化为} \begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}\right)$.

15. 设 A 是 m 阶阵，D 是 n 阶非异阵，B 与 C 分别是 $m\times n$ 与 $n\times m$ 阵。求证：$M\equiv\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 是非异阵 $\Longleftrightarrow A-BD^{-1}C$ 是非异阵；且当 $G\equiv A-BD^{-1}C$ 是非异阵时，成立

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} G^{-1} & -G^{-1}BD^{-1} \\ -D^{-1}CG^{-1} & D^{-1}+D^{-1}CG^{-1}BD^{-1} \end{pmatrix}.$$

上式也称为非异阵求逆的第一降阶定理。

16. 设 A、D 分别是 m 与 n 阶非异阵，B、C 分别是 $m\times n$ 与 $n\times m$ 阵。求证：

(i) 若 $D-CA^{-1}B$ 是非异阵，则 $A-BD^{-1}C$ 也是非异阵，且

$$(A-BD^{-1}C)^{-1}=A^{-1}+A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1};$$

(ii) 若 $A-BD^{-1}C$ 是非异阵，则 $D-CA^{-1}B$ 也是非异阵，且

$$(D-CA^{-1}B^{1})^{-1}=D^{-1}+D^{-1}C(A-BD^{-1}C)^{-1}BD^{-1}.$$

上两式称为非异阵求逆的第二降阶定理。

17. 设 A、B、C、D 是同阶非异阵，且 $D-CA^{-1}B$ 是非异阵。求证：

$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$、$A-BD^{-1}C$、$C-DB^{-1}A$、$B-AC^{-1}D$ 也是非异阵，且

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A-BD^{-1}C)^{-1} & (C-DB^{-1}A)^{-1} \\ (B-AC^{-1}D)^{-1} & (D-CA^{-1}B)^{-1} \end{pmatrix},$$

上式称为非异阵求逆的第三降阶定理。

(提示：利用前两题的结果。)

18. 利用 16 题的结果，证明第 9 题以及 Sherman-Morrison 公式。

第二章　行　列　式

行列式起源于求解线性(代数)方程组．其基础早在十九世纪初已由 Cauchy 所奠定．由于它在理论推导中的重要作用，与矩阵一样，行列式理论早已越出求解线性方程组的范围，而广泛应用于力学、工程数学以及其他科学领域．本章将介绍行列式理论的一些基本内容，并初步把它们应用于线性方程组和矩阵理论方面；介绍有广泛应用的行列式降阶定理以及 Cauchy-Binet 公式．

§1　行列式概念

在初等代数里学过，当二阶行列式

$$D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{1}$$

不等于零时，二元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

有唯一解

$$x_1 = \frac{1}{D_2}\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \quad x_2 = \frac{1}{D_2}\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

当三阶行列式

$$D_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{2}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

不等于零时，三元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

有唯一解

$$x_1 = \frac{1}{D_3}\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad x_2 = \frac{1}{D_3}\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix},$$

$$x_3 = \frac{1}{D_3}\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

由此可见，用行列式表达这些方程组的解，表达式整齐清晰，便于记忆和运用。

由于大量的工程技术问题所归结出来的数学模型往往是含成千上万个未知数和方程的大型线性方程组．因此有必要讨论一般的，即含 n 个未知量、n 个方程的线性方程组，这就是本章所要讨论的主要内容之一．我们把如下方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{3}$$

称为 **n 阶（或 n 元）线性（代数）方程组**．这里，n 是任意正整数，x_j 表示第 j 个未知量，a_{ij} 表示第 i 个方程中 x_j 的系数，b_i 表示第 i 个方程的常数项，i 和 j 是适合 $1 \leqslant i,\ j \leqslant n$ 的任意一对正整数。

从二、三阶行列式对于求解二、三阶线性方程组的作用得到启发，是否可引进所谓 n 阶行列式的概念来讨论 n 阶线性方程组的求解问题？为此，我们先分析一下二阶行列式 D_2 和三阶行列式 D_3 有哪些特点。

在 D_2 和 D_3 的展开式(1)和(2)中可发现以下事实：

(1) D_2 和 D_3 都是一个数，它们分别是 $2! = 2$ 或 $3! = 6$ 项的

代数和，每一项分别是 2 个或 3 个数的乘积，这些乘积中的因子都取之于该行列式中不同的行和不同的列，而且取之于任意不同的行和任意不同的列的 2 个或 3 个数的乘积必定作为它的一个项。

(2) 带正号的项与带负号的项各占一半。如果我们规定：每一项的各个因子 a_{ij} 按这样的顺序写出，使得它们的第一个下标排成自然顺序(正如(1)和(2)式中所写出的那样)，则我们发现，对 D_2 来说，带正号的项所确定的第二个下标的排列为 12，而带负号的项所确定的第二个下标的排列却为 21，它与自然顺序 12 有了一个"颠倒"。同样，对 D_3 来说，每一项的第二个下标的排列，对应于带正号的项是 123，231 和 312，它们的颠倒个数都为偶数(包括零)，对应于带负号的项是 321、213 和 132，它们的颠倒个数都是奇数。这就是说，这种颠倒个数的偶奇性决定了该项所取的正负号。

由上述分析可见，有必要先对若干个数的排列次序作一些讨论。在初等代数中已介绍过由 1，2，…，n 这 n 个数按某种次序所排成的一个数组 $j_1j_2\cdots j_n$ 称为一个 **n 阶排列**，这里 j_1，j_2，…，j_n 是取之于 1，2，…，n 中的 n 个不同的数，j_k 的下标 k 表示数 j_k 在该排列中的第 k 个位置。更一般地，n 个任意数 a_1，a_2，…，a_n 的 n 阶排列为 $a_{j_1}a_{j_2}\cdots a_{j_n}$，这里 $j_1j_2\cdots j_n$ 是 1，2，…，n 这 n 个下标的某个 n 阶排列。易见，n 阶排列的总个数为 $n!=1\cdot2\cdots n$。

我们发现，在 D_2 和 D_3 的展开式(1)和(2)中所出现的项数恰好分别是 2 阶排列和 3 阶排列的总个数。

定义 对 n 阶排列 $j_1j_2\cdots j_n$ 中任意取定的两个数 j_k 和 j_l，如果 j_k 在 j_l 的左边(即前面)而 $j_k>j_l$，则称 j_k 与 j_l 构成了一个**逆序**。一个排列中所有逆序的总个数称为该排列的逆序数，记为 $N(j_1j_2\cdots j_n)$。若 $N(j_1j_2\cdots j_n)$是偶数时，称 $j_1j_2\cdots j_n$ 为**偶排列**，否则，称为**奇排列**。

例如，2413 有 3 个逆序：21，41 和 43，所以它是 4 阶奇排列。

根据这个定义，我们可把二、三阶行列式用统一的形式表示为

$$D_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=\sum_{(j_1j_2)}(-1)^{N(j_1j_2)}a_{1j_1}a_{2j_2},$$

$$D_3=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=\sum_{(j_1j_2j_3)}(-1)^{N(j_1j_2j_3)}a_{1j_1}a_{2j_2}a_{3j_3}.$$

这里求和号

$$\sum_{(j_1,j_2)} \text{ 和 } \sum_{(j_1,j_2,j_3)}$$

分别表示对所有 2! 个不同的 2 阶排列和所有 3! 个不同的 3 阶排列求和。对任一正整数 n，我们引入

定义 设 $A=(a_{ij})_{n\times n}$ 为 n 阶方阵。当 $n=1$ 时，就称 a_{11} 为一阶行列式。当 $n\geqslant 2$ 时，称

$$D_n=\begin{vmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots & \cdots\cdots \\ a_{n1} & a_{n2}\cdots a_{nn} \end{vmatrix}$$

$$=\sum_{(j_1j_2\cdots j_n)}(-1)^{N(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n} \tag{4}$$

为 **n 阶行列式**，也称为方阵 A 的行列式。常记为

$$D_n=|A| \text{ 或 } D_n=\det A \text{ 或 } D_n=|a_{ij}|_{n\times n}.$$

(4)式中的求和号 $\sum\limits_{(j_1j_2\cdots j_n)}$ 表示对所有 $n!$ 个不同的 n 阶排列求和。以后，在不会引起混淆的情况下，往往把它简记为 Σ。

这里自然要产生一个疑问：是否与 D_2 和 D_3 的情形一样，在 D_n 的展开式(4)中，正项与负项也各占一半呢？回答是肯定的。为此，只要证明，在 $n!$ 个 n 阶排列中，奇排列与偶排列的个数相同。如果把排列 $j_1j_2\cdots j_n$ 中某两个数的所在位置对调一下而保持其它数字不动，则称在此排列中施行了一次**对换**。一个排列施行了一次对换后得到了另一个排列。我们有

命题 1 任一排列施行一次对换后，必改变了它的奇偶性。

【证明】 设 $j_1j_2\cdots j_n$ 是 n 阶排列。我们先考虑简单的情形，就是所对调的是相邻的两个数 j_t 和 j_{t+1}。因为 j_t 和 j_{t+1} 与其它数

是否构成逆序不会由于 j_s 与 j_{s+1} 的对调而改变，所以这种对换仅使原排列增加或减少了一个逆序，因而改变了它的奇偶性。

其次，考虑一般情形，此对换对调 j_s 和 j_{s+k+1}，即把

$$j_1\cdots j_s j_{s+1}\cdots\cdots j_{s+k} j_{s+k+1}\cdots\cdots j_n \tag{5}$$

变为

$$j_1\cdots j_{s+k+1} j_{s+1}\cdots\cdots j_{s+k} j_s\cdots\cdots j_n. \tag{6}$$

但这也可以这样得到：从原排列(5)出发，先把 j_{s+k+1} 依次与前面相邻的数对换，如此经过 $k+1$ 次以后，得到排列

$$j_1\cdots j_{s+k+1} j_s j_{s+1}\cdots\cdots j_{s+k} j_{s+k+2}\cdots\cdots j_n.$$

再把 j_s 依次与后面相邻的数对换，经过 k 次后就得到了(6)。这样一共施行了 $2k+1$ 次相邻两数的对换，因而(5)与(6)的奇偶性正好相反。 证毕.

命题 2 在 $n!$ 个不同的 n 阶排列中，奇排列与偶排列各占一半。

【证明】 设 n 阶奇排列有 p 个，n 阶偶排列有 q 个。现在把每个排列中的 1 与 2 施行对换，则 p 个不同的奇排列变为 p 个不同的偶排列，因此 $p\leqslant q$. 同理，$q\leqslant p$. 所以 $p=q$。 证毕。

例 计算下三角形行列式

$$|L|=\begin{vmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ l_{31} & l_{32} & l_{33} & \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots \\ l_{n1} & l_{n2} & l_{n3}\cdots & l_{nn} \end{vmatrix}.$$

由定义易知 $|L|=l_{11}l_{22}l_{33}\cdots\cdots l_{nn}$.

作为本例的特殊情形，第二种初等阵 $D_i(c)$，$c\neq 0$，其行列式为 $|D_i(c)|=c$。对角阵的行列式就是它的所有对角元的乘积，因而单位阵的行列式是 1。第三种初等阵 $T_{ij}(k)$ 的行列式为 1 $(i>j)$。根据本节习题 7 的结果可知，当 $i<j$ 时，$T_{ij}(k)$的行列式也为 1。

习　　题

1. 决定 n 阶排列 $n(n-1)\cdots 2\cdot 1$ 的逆序数，并讨论它的奇偶性。

2. 如果 n 阶排列 $j_1j_2\cdots j_n$ 的逆序数为 k，试求定排列 $j_nj_{n-1}\cdots j_2j_1$ 的逆序数。

3. 设 n 阶排列 $j_1j_2\cdots j_kj_{k+1}\cdots j_n$ 的前 k 个数的每一个都比后 $n-k$ 个数中的任一个要小，求证

$$N(j_1\cdots j_kj_{k+1}\cdots j_n)=N(j_1\cdots j_k)+N(j_{k+1}\cdots j_n).$$

4. 对下面这个 100 阶排列

$$(10)(9)\cdots(1)\cdot(20)(19)\cdots(11)\cdot\cdots\cdots(100)(99)\cdots(91)$$

用最快的速度求出它的逆序数。

5. 在 n 阶排列 $j_1\cdots j_rj_{r+1}\cdots j_n$ 中，若

$$1\leqslant j_1<j_2<\cdots<j_r\leqslant n,\quad 1\leqslant j_{r+1}<j_{r+2}<\cdots<j_n\leqslant n.$$

证明

$$N(j_1\cdots j_rj_{r+1}\cdots j_n)=\sum_{i=1}^{r} j_i-\frac{r(r+1)}{2}.$$

6. 求证

(i)
$$\begin{vmatrix} & & & b_1 \\ & & b_2 & \\ & \iddots & & \\ b_{n-1} & & & \\ b_n & & & \end{vmatrix}=(-1)^{\frac{n(n-1)}{2}}b_1b_2\cdots b_{n-1}b_n;$$

(ii)
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & 0 & 0 & 0 \\ a_{41} & a_{42} & 0 & 0 & 0 \\ a_{51} & a_{52} & 0 & 0 & 0 \end{vmatrix}=0.$$

7. 计算

$$|R|=\begin{vmatrix} r_{11} & r_{12} & r_{13}\cdots r_{1n} \\ & r_{22} & r_{23}\cdots r_{2n} \\ & & \ddots\quad \vdots \\ & & r_{nn} \end{vmatrix}.$$

8. 设 A 是 n 阶复矩阵，$\bar{A}$ 是 A 的共轭矩阵，求证

$$|\bar{A}|=\overline{|A|}.$$

9. 设 $b_{ij}=(-1)^{i+j}a_{ij}$，$i,j=1,2,\cdots,n$。求证 $|b_{ij}|_{n\times n}=|a_{ij}|_{n\times n}$。

§ 2　行列式的性质

虽然在理论上说，总可以由行列式的定义算出任一 n 阶行列式．但当 n 较大时，在实际上，这几乎是不可能的．因为 n 阶行列式共有 $n!$ 项，而每一项都是 n 个数的乘积，所以共需做 $n!(n-1)$ 次乘法．例如，计算一个 20 阶行列式，就要做 $20!19 \doteq 2433\times10^{15}\times 19 \doteq 44127\times10^{15}$ 次乘法①．目前，即便使用高速电子计算机计算，这也是不可能的．因此，必须讨论行列式的基本性质，用来简化行列式的计算．而且，这些基本性质在行列式理论中更为重要．

有时候需要把 n 阶行列式中一般项 $a_{1j_1}a_{2j_2}\cdots a_{nj_n}$ 中因子的排列次序任意调动成

$$a_{1j_1}a_{2j_2}\cdots a_{nj_n}=a_{i_1k_1}a_{i_2k_2}\cdots a_{i_nk_n},$$

这就要求我们找出 $N(j_1j_2\cdots j_n)$、$N(i_1i_2\cdots i_n)$与 $N(k_1k_2\cdots k_n)$ 之间的关系，这就是

引理　如果 $a_{i_1k_1}a_{i_2k_2}\cdots a_{i_nk_n}=a_{1j_1}a_{2j_2}\cdots a_{nj_n}$，则

$$(-1)^{N(i_1i_2\cdots i_n)+N(k_1k_2\cdots k_n)}=(-1)^{N(j_1j_2\cdots j_n)}. \tag{1}$$

【证明】　因为对调 $a_{i_1k_1}a_{i_2k_2}\cdots a_{i_nk_n}$ 中的任意两个因子，对于排列 $i_1i_2\cdots i_n$ 和 $k_1k_2\cdots k_n$ 来说，仅仅是分别施行了一次对换．据§1的命题 1，这并不改变

$$(-1)^{N(i_1i_2\cdots i_n)+N(k_1k_2\cdots k_n)}$$

的数值，因而对于经过了有限次这种对调所得到的 $a_{1j_1}a_{2j_2}\cdots\cdots a_{nj_n}$，必有等式

$$(-1)^{N(i_1i_2\cdots i_n)+N(k_1k_2\cdots k_n)}=(-1)^{N(12\cdots n)+N(j_1j_2\cdots j_n)}.$$

但 $N(12\cdots n)=0$，所以等式(1)成立．　　　　　　证毕．

设 $A=(a_{ij})$是 n 阶阵，A' 是 A的转置阵，则称$|A'|$为$|A|$的转置行列式．

通常所说的行列式的基本性质如下：

---

① 由公式 $n!\approx\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$估算出 $20!\doteq 2433\times10^{15}$，这里 $e\doteq2.71828$．

性质 1　$|A'|=|A|$。即任一矩阵经转置后，行列式不变。

【证明】　设 $A=(a_{ij})_{n\times n}$。记 $A'=(b_{ij})_{n\times n}$，这里 $b_{ij}=a_{ji}$，i，$j=1, 2, \cdots, n$。于是据行列式的定义有

$$|A|=\Sigma(-1)^{N(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots\cdots a_{nj_n},$$

$$\begin{aligned}|A'|&=\Sigma(-1)^{N(i_1i_2\cdots i_n)}b_{1i_1}b_{2i_2}\cdots\cdots b_{ni_n}\\&=\Sigma(-1)^{N(i_1i_2\cdots i_n)}a_{i_11}a_{i_22}\cdots\cdots a_{i_nn}.\end{aligned}$$

由引理不难看出 $|A'|$ 的展开式中的任一项必是 $|A|$ 的展开式中的某一项，且正负号也相同；$|A'|$ 的展开式中任意两项都不相同；$|A'|$ 与 $|A|$ 的展开式中的总项数同为 $n!$。根据这三点就可推得

$$|A'|=|A|.$$

证毕。

性质 1 是非常重要的。它告诉我们，凡是对行列式的"行"所成立的性质对行列式的"列"也一定成立，反过来也是如此。所以，今后我们只要对"列"进行讨论就可以了。

例 1　设 A 是 Hermite 阵，则 $|A|$ 必是实数。

【证明】　因为 $A=\overline{A'}$，所以由性质 1 与 §1 的习题 8 即得

$$|A|=|\overline{A'}|=\overline{|A'|}=\overline{|A|},$$

所以 $|A|$ 是实数。　　证毕。

这里提请读者注意，性质 1 实际上证明了，对 $A=(a_{ij})_{n\times n}$ 有

$$\begin{aligned}|A|&=\Sigma(-1)^{N(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots\cdots a_{nj_n}\\&=\Sigma(-1)^{N(j_1j_2\cdots j_n)}a_{j_11}a_{j_22}\cdots\cdots a_{j_nn}.\end{aligned}\qquad(2)$$

这里 Σ 都是对 $n!$ 个不同的 n 阶排列 $j_1j_2\cdots j_n$ 求和。在证明以下诸性质时常要用到这一等式。

性质 2　若把 n 阶阵 A 按它的列分块

$$A=(\alpha_1, \cdots, \alpha_j, \cdots, \alpha_n),$$

则

$$|\cdots, k\alpha_j, \cdots|=k\cdot|\cdots, \alpha_j, \cdots|,$$

这里 k 是任一数。省略号表示不受影响的其他各列。此性质称为行列式可按列提取公因子，特别有

$$|\alpha_1, \cdots, \alpha_{j-1}, 0, \alpha_{j+1}, \cdots, \alpha_n|=0.$$

【证明】　设 $A=(a_{ij})_{n\times n}$，利用(2)式可得

$$
\begin{aligned}
&|\cdots, k\alpha_j, \cdots| \\
&= \Sigma(-1)^{N(i_1i_2\cdots i_n)} a_{i_11}\cdots(ka_{i_jj})\cdots a_{i_nn} \\
&= k(\Sigma(-1)^{N(i_1i_2\cdots i_n)} a_{i_11}\cdots a_{i_jj}\cdots a_{i_nn}) \\
&= k\cdot|\cdots, \alpha_j, \cdots|.
\end{aligned}
$$

证毕。

例 2

$$
\begin{vmatrix} 2 & 1 & -3 & -1 \\ 7 & 1 & 3 & 2 \\ 3 & 6 & -9 & 1 \\ 4 & 0 & 24 & 8 \end{vmatrix} = 12\cdot\begin{vmatrix} 2 & 1 & -1 & -1 \\ 7 & 1 & 1 & 2 \\ 3 & 6 & -3 & 1 \\ 1 & 0 & 2 & 2 \end{vmatrix}.
$$

性质 3 如果A的第j列α_j为两个列向量之和：$\alpha_j=\beta+\gamma$，则

$$
|\cdots, \alpha_j, \cdots| = |\cdots, \overset{j}{\beta}, \cdots| + |\cdots, \overset{j}{\gamma}, \cdots|.
$$

【证明】 设$A=(a_{ij})_{n\times n}=(\alpha_1, \cdots, \alpha_j, \cdots, \alpha_n)$。记

$$
\beta=(b_1, b_2, \cdots, b_n)', \gamma=(c_1, c_2, \cdots, c_n)',
$$

则

$$
\alpha_j=(b_1+c_1, b_2+c_2, \cdots, b_n+c_n)',
$$

于是利用(2)式就有

$$
\begin{aligned}
|A| &= |\alpha_1, \cdots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \cdots, \alpha_n| \\
&= \Sigma(-1)^{N(i_1i_2\cdots i_n)} a_{i_11}\cdots a_{i_{j-1},j-1}\cdot(b_j+c_j)\cdot a_{i_{j+1},j+1}\cdots a_{i_nn} \\
&= \Sigma(-1)^{N(i_1i_2\cdots i_n)} a_{i_11}\cdots a_{i_{j-1},j-1}\cdot b_j\cdot a_{i_{j+1},j+1}\cdots a_{i_nn} \\
&\quad + \Sigma(-1)^{N(i_1i_2\cdots i_n)} a_{i_11}\cdots a_{i_{j-1},j-1}\cdot c_j\cdot a_{i_{j+1},j+1}\cdots a_{i_nn} \\
&= |\alpha_1, \cdots, \alpha_{j-1}, \beta, \alpha_{j+1}, \cdots, \alpha_n| \\
&\quad + |\alpha_1, \cdots, \alpha_{j-1}, \gamma, \alpha_{j+1}, \cdots, \alpha_n|.
\end{aligned}
$$

证毕。

例 3 连续应用性质 3 可得

$$
\begin{vmatrix} a_1+a_2 & b_1+b_2 \\ c_1+c_2 & d_1+d_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix} + \begin{vmatrix} a_1 & b_2 \\ c_1 & d_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_1 \\ c_2 & d_1 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_2 & d_2 \end{vmatrix},
$$

但是，一般说来

$$
\begin{vmatrix} a_1+a_2 & b_1+b_2 \\ c_1+c_2 & d_1+d_2 \end{vmatrix} \neq \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_2 & d_2 \end{vmatrix}.
$$

性质 4 若A中某两列相等，则$|A|=0$。

【证明】 设

$$A=(a_{ij})_{n\times n}=(\alpha_1,\ \cdots,\ \overset{j}{\alpha_j},\ \cdots,\ \overset{k}{\alpha_j},\ \cdots,\ \alpha_n),$$

其中第 k 列与第 j 列同为 $\alpha_j=(a_{1j},\ a_{2j},\ \cdots,\ a_{nj})'$。于是利用(2)式就有

$$|A|=\Sigma(-1)^{N(i_1i_2\cdots i_n)}a_{i_11}\cdots a_{i_jj}\cdots a_{i_kk}\cdots a_{i_nn},$$

对于其中任一项

$$(-1)^{N(i_1\cdots i_j\cdots i_k\cdots i_n)}a_{i_11}\cdots a_{i_jj}\cdots a_{i_kk}\cdots a_{i_nn}$$

来说，由行列式定义知必有另一项

$$(-1)^{N(i_1\cdots i_k\cdots i_j\cdots i_n)}a_{i_11}\cdots a_{i_kj}\cdots a_{i_jk}\cdots a_{i_nn}$$

与它配对，它们的差别仅在于第 j 个因子与第 k 个因子的取法上。但由条件知 $a_{i_kk}=a_{i_kj}$，$a_{i_jk}=a_{i_jj}$，而 n 阶排列 $i_1\cdots i_j\cdots i_k\cdots i_n$ 与 $i_1\cdots i_k\cdots i_j\cdots i_n$ 的奇偶性正好相反，所以上述相应两项之和为零。由于取定的一个对换把 $n!$ 个 n 阶排列按上述方式两两配对，所以 $|A|$ 中的 $n!$ 项两两相消，即得 $|A|=0$。 证毕。

性质 5 若 A 中某两列成比例，则 $|A|=0$。

【证明】 假设 $\alpha_k=s\cdot\alpha_j$，s 是某个数。利用性质 2 和 4 即得

$$\begin{aligned}|A|&=|\alpha_1,\ \cdots,\ \alpha_j,\ \cdots,\ s\cdot\alpha_j,\ \cdots,\ \alpha_n|\\&=s\cdot|\alpha_1,\ \cdots,\ \alpha_j,\ \cdots,\ \alpha_j,\ \cdots,\ \alpha_n|=s\cdot 0=0.\end{aligned}$$ 证毕。

性质 6 $|\cdots,\ \alpha_j,\ \cdots,\ \alpha_k,\ \cdots|=|\cdots,\ \alpha_j,\ \cdots,\ \alpha_k+s\cdot\alpha_j,\ \cdots|$。即把行列式的某一列的 s 倍加到另一列上去，行列式不变。

这是性质 3 和性质 5 的直接推论。

性质 7 $|\cdots,\ \overset{j}{\alpha_j},\ \cdots,\ \overset{k}{\alpha_k},\ \cdots|=-|\cdots,\ \overset{j}{\alpha_k},\ \cdots,\ \overset{k}{\alpha_j},\ \cdots|$。即互换两列，行列式改号。

【证明】 由性质 4 和性质 3 即得

$$\begin{aligned}0&=|\cdots,\ \alpha_j+\alpha_k,\ \cdots,\ \alpha_j+\alpha_k,\ \cdots|\\&=|\cdots,\ \alpha_j,\ \cdots,\ \alpha_k,\ \cdots|+|\cdots,\ \alpha_j,\ \cdots,\ \alpha_j,\ \cdots|\\&\quad+|\cdots,\ \alpha_k,\ \cdots,\ \alpha_k,\ \cdots|+|\cdots,\ \alpha_k,\ \cdots,\ \alpha_j,\ \cdots|\\&=|\cdots,\ \alpha_j,\ \cdots,\ \alpha_k,\ \cdots|+|\cdots,\ \alpha_k,\ \cdots,\ \alpha_j,\ \cdots|,\end{aligned}$$

这里每个行列式中所写出来的两个列向量都是位于第 j 和第 k 两列。 证毕。

例 4 第一种初等阵 P_{jk} 的行列式等于 -1。

【证明】 $$|P_{jk}|=|e_1,\cdots,\overset{j}{e_k},\cdots,\overset{k}{e_j},\cdots,e_n|$$
$$=-|e_1,\cdots,e_j,\cdots,e_k,\cdots,e_n|$$
$$=-|I_n|=-1.$$ 证毕。

最后，我们举几个例子说明怎样利用行列式性质简化行列式的计算过程。

例 5 当 $n\geqslant 3$ 时有

$$\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & 1+x_1y_3\cdots\cdots 1+x_1y_n \\ 1+x_2y_1 & 1+x_2y_2 & 1+x_2y_3\cdots\cdots 1+x_2y_n \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ 1+x_ny_1 & 1+x_ny_2 & 1+x_ny_3\cdots\cdots 1+x_ny_n \end{vmatrix}$$
$$=\begin{vmatrix} x_1(y_1-y_n) & x_1(y_2-y_n) & 1+x_1y_3\cdots\cdots 1+x_1y_n \\ x_2(y_1-y_n) & x_2(y_2-y_n) & 1+x_2y_3\cdots\cdots 1+x_2y_n \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ x_n(y_1-y_n) & x_n(y_2-y_n) & 1+x_ny_3\cdots\cdots 1+x_ny_n \end{vmatrix}$$
$$=0,$$

这里，是在第 1 列和第 2 列中分别减去最后一列中相应元素而得到了前二列成比例的行列式，因而其值为零。

一般说来，计算数字行列式的基本想法是，应用行列式诸性质把它化成上或下三角形行列式，从而求出其值。

例 6 计算行列式

$$D=\begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 8 & -7 & -10 \end{vmatrix}.$$

【解】 把 D 的第 1 行与第 2 行对调得到

$$D=-\begin{vmatrix} 1 & -9 & 13 & 7 \\ -2 & 5 & -1 & 3 \\ 3 & -1 & 5 & -5 \\ 2 & 8 & -7 & -10 \end{vmatrix},$$

而

$$\begin{array}{l}(-2)(-3)(2)\end{array}\begin{vmatrix}1 & -9 & 13 & 7\\ -2 & 5 & -1 & 3\\ 3 & -1 & 5 & -5\\ 2 & 8 & -7 & -10\end{vmatrix}$$

$$=\begin{array}{l}(2)\ (2)\end{array}\begin{vmatrix}1 & -9 & 13 & 7\\ 0 & -13 & 25 & 17\\ 0 & 26 & -34 & -26\\ 0 & 26 & -33 & -24\end{vmatrix}$$

$$=\left(-\frac{17}{16}\right)\begin{vmatrix}1 & -9 & 13 & 7\\ 0 & -13 & 25 & 17\\ 0 & 0 & 16 & 8\\ 0 & 0 & 17 & 10\end{vmatrix}$$

$$=\begin{vmatrix}1 & -9 & 13 & 7\\ 0 & -13 & 25 & 17\\ 0 & 0 & 16 & 8\\ 0 & 0 & 0 & \frac{3}{2}\end{vmatrix}$$

$$=(-13)\cdot 16\cdot\frac{3}{2}=-312,$$

所以 $D=312$。

这里，行列式左边带箭头的折线表示所施行的初等行变换，与第一章中矩阵的初等变换意义相同。

例 7 设 $a_i\neq 0$, $i=1, 2, \cdots, n$. 计算

$$D=\begin{vmatrix}a_1 & 1 & \cdots & 1\\ 1 & a_2 & & \\ \vdots & & \ddots & \\ 1 & & & a_n\end{vmatrix}.$$

【解】 把 D 的第 2 列的 $-\frac{1}{a_2}$ 倍，第 3 列的 $-\frac{1}{a_3}$ 倍，……，第 n 列的 $-\frac{1}{a_n}$ 倍统统加到第 1 列上，得到

$$D=\begin{vmatrix} a_1-\sum\limits_{i=2}^{n}\frac{1}{a_i} & 1 & \cdots & 1 \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{vmatrix}$$

$$=a_2\cdots a_n\left(a_1-\sum_{i=2}^{n}\frac{1}{a_i}\right).$$

习　　题

1．计算下列行列式

(i) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$； (ii) $\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix}$；

(iii) $\begin{vmatrix} a+b & b+c & c+a \\ b+c & c+a & a+b \\ c+a & a+b & b+c \end{vmatrix}$； (iv) $\begin{vmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{vmatrix}$.

2．求证 $\begin{vmatrix} a_1+b_1 & b_1+c_1 & c_1+a_1 \\ a_2+b_2 & b_2+c_2 & c_2+a_2 \\ a_3+b_3 & b_3+c_3 & c_3+a_3 \end{vmatrix}=2\cdot\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$.

3．试由 204，527 和 255 这三个三位数都能被 17 整除这一事实说明三阶行列式

$$D=\begin{vmatrix} 2 & 0 & 4 \\ 5 & 2 & 7 \\ 2 & 5 & 5 \end{vmatrix}$$

必能被 17 整除而不需要求出 D 的值.

4．证明奇数阶反对称阵的行列式必为零.

5．计算以下 n 阶行列式：

(i) $D=\begin{vmatrix} x & a & a & \cdots & a & a \\ -a & x & a & \cdots & a & a \\ -a & -a & x & \cdots & a & a \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -a & -a & -a & \cdots & x & a \\ -a & -a & -a & \cdots & -a & x \end{vmatrix}$；

(ii) $D=\begin{vmatrix} b_1 & x & x\cdots\cdots & x & x \\ x & b_2 & x\cdots\cdots & x & x \\ x & x & b_3\cdots\cdots & x & x \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x & x & x\cdots\cdots & b_{n-1} & x \\ x & x & x\cdots\cdots & x & b_n \end{vmatrix}$.

6. 应用 Vandermonde 行列式计算如下行列式

(i) $\begin{vmatrix} x_1^{n-1} & x_1^{n-2}\cdots\cdots x_1 & 1 \\ x_2^{n-1} & x_2^{n-2}\cdots\cdots x_2 & 1 \\ \cdots\cdots & \cdots\cdots & \cdots \\ x_n^{n-1} & x_n^{n-2}\cdots\cdots x_n & 1 \end{vmatrix}$,

(ii) $\begin{vmatrix} a_1^{n-1} & a_1^{n-2}b_1\cdots\cdots a_1b_1^{n-2} & b_1^{n-1} \\ a_2^{n-1} & a_2^{n-2}b_2\cdots\cdots a_2b_2^{n-2} & b_2^{n-1} \\ \cdots\cdots & \cdots\cdots & \cdots \\ a_n^{n-1} & a_n^{n-2}b_n\cdots\cdots a_nb_n^{n-2} & b_n^{n-1} \end{vmatrix}$,

(iii) $\begin{vmatrix} 1 & x_1+1 & x_1^2+x_1 & \cdots & x_1^{n-1}+x_1^{n-2} \\ 1 & x_2+1 & x_2^2+x_2 & \cdots & x_2^{n-1}+x_2^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n+1 & x_n^2+x_n & & x_n^{n-1}+x_n^{n-2} \end{vmatrix}$,

(iv) $\begin{vmatrix} \frac{x_1}{x_1-1} & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \frac{x_2}{x_2-1} & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{x_n}{x_n-1} & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}$,

(V) $\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} & x_2^n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} & x_n^n \end{vmatrix}$.

§3 行列式的乘法规则

在 §1 末和 §2 的例 4 中已经证明了三种初等阵的行列式为

$$|P_{ij}|=-1,\ |C_i(c)|=c,\ |T_{ij}(k)|=1.$$

据此可得以下引理

引理 设P是n阶初等阵，A是任意n阶阵，则

$$|AP|=|A|\cdot|P|,\ |PA|=|P|\cdot|A|.$$

【证明】 我们先证明第一个等式。事实上，这个P无非是三种初等阵P_{ij}，$D_i(c)$，$(c\neq0)$和$T_{ij}(k)$之一，据行列式性质7，2和6，依次有

$$|AP_{ij}|=-|A|=|A|\cdot|P_{ij}|.$$
$$|AD_i(c)|=c\cdot|A|=|A|\cdot|D_i(c)|.$$
$$|AT_{ij}(k)|=|A|=|A|\cdot|T_{ij}(k)|.$$

据行列式性质1可从第一个等式推得第二个等式

$$|PA|=|(PA)'|=|A'P'|=|A'|\cdot|P'|$$
$$=|A|\cdot|P|=|P|\cdot|A|.$$

这里用到初等阵的转置仍是初等阵，所以对P'可应用第一个等式。 证毕。

定理3.1 （行列式乘法规则）设A和B是任意两个n阶阵，则$|AB|=|A|\cdot|B|$。

【证明】 区别以下两种情形。

1．若B是非异阵。由第一章定理5.2可知B可表为有限个初等阵之积$B=P_1P_2\cdots P_s$。于是连续应用引理就得

$$|AB|=|AP_1P_2\cdots P_s|=|A|\cdot|P_1|\cdot|P_2|\cdots|P_s|$$
$$=|A|\cdot|P_1P_2\cdots P_s|=|A|\cdot|B|.$$

仿此可证$|BA|=|B|\cdot|A|$。这说明对非异阵B必有

$$|AB|=|BA|=|A|\cdot|B|. \tag{1}$$

2．若B是奇异阵。由第一章定理5.1′的推论1知存在n阶非异阵P和Q使得

$$B=P\begin{pmatrix}I_s & 0\\ 0 & 0\end{pmatrix}Q,$$

且必有$s<n$。因为否则的话，B将是非异阵，这与假设矛盾。应用(1)式可得

$$|B| = |P| \cdot \begin{vmatrix} I_s & 0 \\ 0 & 0 \end{vmatrix} \cdot |Q| = 0$$

和

$$|AB| = \left| AP\begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} Q \right| = \left| AP\begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} \right| \cdot |Q|.$$

但是，由 $s<n$ 知 $AP\begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}$的最后一列必为零，其行列式为零，所以 $|AB|=0$。此时，$|AB|=|A|\cdot|B|$也成立。 证毕。

定理 3.2 设 A 是 n 阶阵，则 A 是非异阵的充分必要条件是 $|A|$不等于零。

【证明】 设 A 是非异阵，则存在 B 使 $|AB|=I_n$。由行列式乘法规则得 $|A|\cdot|B|=1$，故必有 $|A|\neq 0$。在定理 3.1 的第二部份证明中已经证得：如 A 是奇异阵，则必有 $|A|=0$，所以，若 $|A|\neq 0$，则 A 必为非异阵。 证毕。

例 1 上(下)三角阵是非异阵当且仅当它的所有主对角元都是非零数。

例 2 正交阵的行列式等于 1 或 -1。

事实上，由 $AA'=I_n$ 即得

$$|A|^2 = |A|\cdot|A'| = |AA'| = |I_n| = 1,$$

所以 $|A|=1$ 或 -1。

例 3 设 A 和 B 都是 n 阶正交阵，且 $|A|+|B|=0$，则 $|A+B|=0$。

【证明】 由条件知 $AA'=B'B=I_n$，所以由行列式乘法规则及行列式性质 1 得

$$\begin{aligned} |A+B| &= |AB'B+AA'B| \\ &= |A|\cdot|B'+A'|\cdot|B| = |A|\cdot|B|\cdot|A+B|. \end{aligned}$$

把 $|B|=-|A|$代入并利用 $|A|^2=1$ 得

$$|A+B| = -|A|^2\cdot|A+B| = -|A+B|.$$

因为 $|A+B|$是数，所以必有 $|A+B|=0$。 证毕。

例 4 用行列式乘法规则证明 §2 中例 5 的结论。

考虑矩阵等式

$$A=\begin{pmatrix}1+x_1y_1 & 1+x_1y_2\cdots\cdots 1+x_1y_n\\ 1+x_2y_1 & 1+x_2y_2\cdots\cdots 1+x_2y_n\\ \cdots\cdots & \cdots\cdots\cdots\cdots\\ 1+x_ny_1 & 1+x_ny_2\cdots\cdots 1+x_ny_n\end{pmatrix}$$

$$=\begin{pmatrix}1 & x_1 & \\ 1 & x_2 & \\ \vdots & \vdots & 0\\ 1 & x_n & \end{pmatrix}\begin{pmatrix}1 & 1\cdots\cdots 1\\ y_1 & y_2\cdots\cdots y_n\\ & 0\end{pmatrix}$$

因为 $n\geqslant 3$，右端两个矩阵的行列式都为零，所以用行列式乘法规则即证得 $|A|=0$。 证毕。

习　　题

1. (i) 求证

$$\begin{vmatrix}a_1 & a_2 & a_3\\ a_3 & a_1 & a_2\\ a_2 & a_3 & a_1\end{vmatrix}=a_1^3+a_2^3+a_3^3-3a_1a_2a_3;$$

(ii) 从而应用行列式乘法规则证明

$$(a_1^3+a_2^3+a_3^3-3a_1a_2a_3)(b_1^3+b_2^3+b_3^3-3b_1b_2b_3)$$
$$=(c_1^3+c_2^3+c_3^3-3c_1c_2c_3),$$

这里 c_1, c_2, c_3 是由 $a_1, a_2, a_3, b_1, b_2, b_3$ 经过加、减、乘三种运算得出的数。

2. 求证下面的三阶阵必是奇异阵

$$\begin{pmatrix}\cos 2x & \cos(x+y) & \cos(x+z)\\ \cos(y+x) & \cos 2y & \cos(y+z)\\ \cos(z+x) & \cos(z+y) & \cos 2z\end{pmatrix}.$$

3. 求证

(i) 如果A是对合阵，即 $A^2=I$，则$|A|=\pm 1$；

(ii) 如果A是 s 次幂零阵，即 $A\neq 0$，且存在自然数 s 使 $A^s=0$，则 $|A|=0$；

(iii) 如果A是非异阵，则$|A^{-1}|=|A|^{-1}$。

4. 设A和B都是 n 阶对合阵，且$|A|+|B|=0$，求证$|A+B|=0$。

5. 记

$$S=\begin{pmatrix} s_0 & s_1 & s_2\cdots\cdots s_{n-1} \\ s_1 & s_2 & s_3\cdots\cdots s_n \\ s_2 & s_3 & s_4\cdots\cdots s_{n+1} \\ \cdots & \cdots & \cdots \\ s_{n-1} & s_n & s_{n+1}\cdots s_{2n-2} \end{pmatrix}$$

其中，$s_k=x_1^k+x_2^k+\cdots+x_n^k$，$k$ 是非负整数，求证 $|S|\geqslant 0$，而不等号成立的充分必要条件是 $x_i\neq x_j$，$i\neq j$，$i, j=1, 2, \cdots, n$。

6. 设 A 是 n 阶非异对称阵，B 是 n 阶反对称阵，求证 $|A+\sqrt{-1}B|^2=|A|^2\cdot|I_n+(A^{-1}B)^2|$。

（提示：由假设先证 $A+\sqrt{-1}B$ 是 Hermite 阵，然后由 Hermite 阵的行列式必是实数这一性质，把 $|A+\sqrt{-1}B|^2$ 写成 $|A+\sqrt{-1}B|\cdot|A-\sqrt{-1}B|$。）

§4 行列式的展开、Cramer 法则

一、n 阶行列式的展开式

本段将把 n 阶行列式写成 n 个 $n-1$ 阶行列式之和，这种降阶的方法被称为行列式按它的一列（或一行）展开。它主要用于理论推导中，有时也可用来计算行列式。为此，先引进

定义 在 n 阶行列式 $|A|=|a_{ij}|_{n\times n}$ 中划去第 i 行元素和第 j 列元素后所剩下的 $(n-1)^2$ 个元素按原来的相对位置所排成的 $n-1$ 阶行列式

$$M_{ij}=\left|A\begin{pmatrix}1\cdots\cdots(i-1)(i+1)\cdots\cdots n\\1\cdots\cdots(j-1)(j+1)\cdots\cdots n\end{pmatrix}\right|$$

称为元素 a_{ij}（在 $|A|$ 中）的**余子式**，而

$$A_{ij}=(-1)^{i+j}M_{ij}$$

称为 a_{ij}（在 $|A|$ 中）的**代数余子式**，这里 $1\leqslant i,\ j\leqslant n$。

例如，

$$|A|=\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

中 a_{23} 的余子式和代数余子式分别是

$$M_{23}=\left|A\begin{pmatrix}1 & 3 & 4\\1 & 2 & 4\end{pmatrix}\right|=\begin{vmatrix}a_{11} & a_{12} & a_{14}\\a_{31} & a_{32} & a_{34}\\a_{41} & a_{42} & a_{44}\end{vmatrix},$$

$$A_{23}=(-1)^{2+3}M_{23}=-M_{23}.$$

定理 4·1 设 $|A|=|a_{ij}|_{n\times n}$ 为 n 阶行列式，A_{ij} 是 a_{ij} 的代数余子式，则

$$\sum_{i=1}^{n} a_{ij}A_{ik}=\delta_{jk}|A|, \tag{1}$$

$$\sum_{i=1}^{n} a_{ji}A_{ki}=\delta_{jk}|A|. \tag{2}$$

这里 δ_{jk} 是 Kronecker 符号，$j,k=1,2,\cdots n$。

【证明】 先证明(1)式。设

$$A=(a_{ij})_{n\times n}=(\alpha_1,\alpha_2,\cdots,\alpha_n),$$

区别以下两种情形：

1．$j=k$。此时，要证明对任一 $1\leqslant j\leqslant n$，有

$$|A|=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots\cdots+a_{nj}A_{nj}. \tag{3}$$

称(3)式为 $|A|$ 按它的第 j 列的展开式。我们把证明过程分成以下三步进行。

(i) 如果

$$|A|=\begin{vmatrix}A\begin{pmatrix}1 & 2\cdots n-1\\1 & 2\cdots n-1\end{pmatrix} & 0\\ * & a_{nn}\end{vmatrix}$$

$$=|\alpha_1,\ \cdots,\ \alpha_{n-1},\ a_{nn}e_n|,$$

则由 §2 的(2)式有

$$\begin{aligned}|A|&=\sum(-1)^{N(j_1j_2\cdots j_n)}a_{j_11}a_{j_22}\cdots\cdots a_{j_{n-1}n-1}a_{j_nn}\\&=\sum(-1)^{N(j_1j_2\cdots j_{n-1}n)}a_{j_11}a_{j_22}\cdots\cdots a_{j_{n-1}n-1}a_{nn}\\&=(\sum(-1)^{N(j_1j_2\cdots j_{n-1})}a_{j_11}a_{j_22}\cdots\cdots a_{j_{n-1}n-1})a_{nn}\\&=\left|A\begin{pmatrix}1 & 2\cdots n-1\\1 & 2\cdots n-1\end{pmatrix}\right|\cdot a_{nn}=a_{nn}A_{nn}.\end{aligned}$$

(ii) 如果

$$|A|=|\alpha_1, \cdots, \alpha_{j-1}, a_{ij}e_i, \alpha_{j+1}, \cdots, \alpha_n|,$$

即A的第 j 列 $\alpha_j=(0, \cdots, 0, a_{ij}, 0, \cdots, 0)'$ 中第 i 个元素为a_{ij}。先把 a_{ij} 经过$n-i$次相邻行的对调调到第(n, j)位置上，再经过$n-j$次相邻列的对调调到第(n, n)位置，即右下角上，这样，连续应用行列式性质 7 就得到

$$|A|=(-1)^{n-i+n-j}\begin{vmatrix} A\begin{pmatrix}1\cdots(i-1)\ (i+1)\cdots n\\ 1\cdots(j-1)(j+1)\cdots n\end{pmatrix} & 0 \\ * & a_{ij}\end{vmatrix}$$

$$=(-1)^{i+j}a_{ij}M_{ij}=a_{ij}A_{ij}.$$

(iii) 一般情形，把A的第 j 列拆成

$$\alpha_j=a_{1j}e_1+a_{2j}e_2+\cdots+a_{nj}e_n=\sum_{i=1}^n a_{ij}e_i,$$

这里Σ是对 n 个列向量 $a_{ij}e_i$ 求和，也就是对其相应分量求和。此时，由行列式性质 3 就得到

$$|A|=|\alpha_1, \cdots, \alpha_{j-1}, \sum_{i=1}^n a_{ij}e_i, \alpha_{j+1}, \cdots, \alpha_n|$$

$$=\sum_{i=1}^n |\alpha_1, \cdots\alpha_{j-1}, a_{ij}e_i, \alpha_{j+1}, \cdots, \alpha_n|$$

$$=\sum_{i=1}^n a_{ij}A_{ij}.$$

于是(3)式成立。

2. $j\neq k$. 此时要证明

$$a_{1j}A_{1k}+a_{2j}A_{2k}+\cdots+a_{nj}A_{nk}=0. \tag{4}$$

为此，可构造如下的第 j 列和第 k 列是相同的 n 阶行列式(因而必为零)：

$$|B|=|\alpha_1, \cdots, \alpha_j, \cdots, \alpha_j, \cdots, \alpha_n|=0.$$

把$|B|$按第 k 列展开就得(4)式。

合并(3)和(4)两式就得到(1)式，再据行列式性质 1，由(1)式就可推得(2)式。　　　　证毕。

在(2)式中取 $j=k$ 所得到的等式

$$|A|=a_{j1}A_{j1}+a_{j2}A_{j2}+\cdots+a_{jn}A_{jn} \tag{5}$$

称为$|A|$按它的第i行的展开式。

例1 计算如下三对角行列式

$$D_n=\begin{vmatrix} a+b & ab & & & \\ 1 & a+b & ab & & \\ & 1 & a+b & \ddots & \\ & & \ddots & a+b & ab \\ & & & 1 & a+b \end{vmatrix}.$$

【解】 把D_n按第1列展开得

$$D_n=(a+b)D_{n-1}-\begin{vmatrix} ab & & & & & \\ 1 & a+b & ab & & & \\ & 1 & a+b & ab & & \\ & & 1 & \ddots & \ddots & \\ & & & \ddots & a+b & ab \\ & & & & 1 & a+b \end{vmatrix},$$

再把上式右边第2个$n-1$阶行列式按它的第1行展开得

$$D_n=(a+b)D_{n-1}-abD_{n-2}. \tag{6}$$

先把(6)式改写为

$$D_n-bD_{n-1}=a(D_{n-1}-bD_{n-2}).$$

连续运用这个关于D_n-bD_{n-1}的递推关系式得

$$D_n-bD_{n-1}=a^{n-2}(D_2-bD_1). \tag{7}$$

再把(6)式改写为

$$D_n-aD_{n-1}=b(D_{n-1}-aD_{n-2}),$$

同理可得

$$D_n-aD_{n-1}=b^{n-2}(D_2-aD_1). \tag{8}$$

由直接计算可知

$$D_1=a+b,\quad D_2=a^2+b^2+ab,$$

代入(7)与(8)式得到

$$D_n-bD_{n-1}=a^n,\ D_n-aD_{n-1}=b^n. \tag{9}$$

当$a\neq b$时，由上式可解得

$$D_n=\frac{1}{a-b}(a^{n+1}-b^{n+1})。$$

当 $a=b$ 时,(9)式化为 $D_n=aD_{n-1}+a^n$。连续运用此递推关系式可得

$$D_n=a^{n-1}D_1+(n-1)a^n=(n+1)a^n。$$

一般地,如果对形如

$$D_n=\begin{vmatrix} \alpha & \beta & & & & \\ \gamma & \alpha & \beta & & & \\ & \gamma & \alpha & \beta & & \\ & & \ddots & \ddots & \ddots & \\ & & & & \alpha & \beta \\ & & & & \gamma & \alpha \end{vmatrix}$$

的三对角行列式,已求得了如下递推关系式

$$D_n=\alpha D_{n-1}-\beta\gamma D_{n-2},$$

则取 a 与 b 是 $x^2-\alpha x+\beta\gamma=0$ 的两个根,则必有 $\alpha=a+b,\beta\gamma=ab$,所以

$$D_n=(a+b)D_{n-1}-abD_{n-2}。$$

于是同例 1 中所证明的一样有

$$D_n=\begin{cases} \dfrac{a^{n+1}-b^{n+1}}{a-b}, & 若\ a\neq b; \\ (n+1)a^n, & 若\ a=b。\end{cases}$$

注意:使用递推关系式求行列式时,必须确定此关系式所适用的范围。例如,对

$$D_n=\begin{vmatrix} a_1 & b_1 & & & & & \\ c_1 & a_2 & b_2 & & & & \\ & c_2 & \alpha & \beta & & & \\ & & \gamma & \alpha & \beta & & \\ & & & \gamma & \alpha & \beta & \\ & & & & \ddots & \ddots & \ddots \\ & & & & & \alpha & \beta \\ & & & & & \gamma & \alpha \end{vmatrix}$$

按最后一行展开后可求得

$$D_n=\alpha D_{n-1}-\beta\gamma D_{n-2}=(a+b)D_{n-1}-abD_{n-2},$$

这里 a 与 b 是 $x^2-\alpha x+\beta\gamma=0$ 的根，但这仅对 $n\geqslant5$ 才成立。因此，仅能推得

$$D_n-bD_{n-1}=a^{n-5}(D_5-bD_4),$$

$$D_n-aD_{n-1}=b^{n-5}(D_5-aD_4).$$

此时，需具体求出 D_3、D_4 和 $D_5=\alpha D_4-\beta\gamma D_3$，才能求出 $D_n, n\geqslant5$。

例 2 设 A 是 n 阶阵，B 是 m 阶阵，C 是 $m\times n$ 阵，则

$$\begin{vmatrix} A & O \\ C & B \end{vmatrix}=|A|\cdot|B|.$$

【证明】 由行列式乘法规则得

$$\begin{vmatrix} A & O \\ C & B \end{vmatrix}=\begin{vmatrix} A & O \\ O & I_m \end{vmatrix}\cdot\begin{vmatrix} I_n & O \\ C & B \end{vmatrix},$$

把上式右端第一个行列式依次按它的第 $n+m$ 行，第 $n+m-1$ 行，……，第 $n+1$ 展开，使得

$$\begin{vmatrix} A & O \\ O & I_m \end{vmatrix}=|A|.$$

同理可证

$$\begin{vmatrix} I_n & O \\ C & B \end{vmatrix}=|B|.$$

于是结论成立。 证毕。

这个例子说明，若能用初等变换把某 n 阶行列式化为

$$\begin{vmatrix} A & O \\ C & B \end{vmatrix} \text{或} \begin{vmatrix} A & C \\ O & B \end{vmatrix},$$

这里 A 和 B 都是阶数小于 n 的方阵，则原行列式就是 $|A|\cdot|B|$。

例 3 称 n 阶行列式

$$V_n=|x_i^{j-1}|_{n\times n}=\begin{vmatrix} 1 & x_1 & x_1^2\cdots\cdots x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2\cdots\cdots x_2^{n-2} & x_2^{n-1} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ 1 & x_{n-1} & x_{n-1}^2\cdots\cdots x_{n-1}^{n-2} & x_{n-1}^{n-1} \\ 1 & x_n & x_n^2\cdots\cdots x_n^{n-2} & x_n^{n-1} \end{vmatrix}$$

为 **Vandermonde 行列式**。今来计算 V_n。

从 V_n 的最后第 2 列开始，依次把前一列的$(-x_n)$倍加到后一列上去，得到

$$V_n=\begin{vmatrix} 1 & x_1-x_n & x_1(x_1-x_n)\cdots\cdots x_1^{n-3}(x_1-x_n) & x_1^{n-2}(x_1-x_n) \\ 1 & x_2-x_n & x_2(x_2-x_n)\cdots\cdots x_2^{n-3}(x_2-x_n) & x_2^{n-2}(x_2-x_n) \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n-1}-x_n & x_{n-1}(x_{n-1}-x_n)\cdots x_{n-1}^{n-3}(x_{n-1}-x_n) & x_{n-1}^{n-2}(x_{n-1}-x_n) \\ 1 & 0 & 0 \quad \cdots \quad 0 & 0 \end{vmatrix}$$

$$=(-1)^{n+1}(x_1-x_n)(x_2-x_n)\cdots\cdots(x_{n-1}-x_n)V_{n-1}$$

$$=(x_n-x_1)(x_n-x_2)\cdots\cdots(x_n-x_{n-1})V_{n-1}.$$

这就是要找的递推公式。于是有如下诸等式

$$V_n=(x_n-x_1)(x_n-x_2)\cdots\cdots(x_n-x_{n-2})(x_n-x_{n-1})V_{n-1},$$

$$V_{n-1}=(x_{n-1}-x_1)(x_{n-1}-x_2)\cdots\cdots(x_{n-1}-x_{n-2})V_{n-2},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$V_3=(x_3-x_1)(x_3-x_2)V_2,$$

$$V_2=(x_2-x_1).$$

逐次往上代入前一等式，即得

$$\begin{aligned} V_n=&(x_n-x_1)(x_n-x_2)\cdots\cdots(x_n-x_{n-2})(x_n-x_{n-1}) \\ &\times(x_{n-1}-x_1)(x_{n-1}-x_2)\cdots\cdots(x_{n-1}-x_{n-2}) \\ &\times\cdots\cdots\cdots\cdots \\ &\times(x_3-x_1)(x_3-x_2)\times(x_2-x_1) \\ =&\prod_{1\leqslant j<i\leqslant n}(x_i-x_j). \end{aligned}$$

这里连乘号 $\prod\limits_{1\leqslant j<i\leqslant n}$ 表示对所有满足 $1\leqslant j<i\leqslant n$ 的因子 x_i-x_j 求其乘积①。

---

① 求积记号定义如下：

$$\prod_{i=1}^{n}a_i=a_1a_2\cdots a_n,$$

$$\prod_{i,j=1}^{n}a_{ij}=\prod_{i=1}^{n}\prod_{j=1}^{n}a_{ij}=\prod_{i=1}^{n}\left(\prod_{j=1}^{n}a_{ij}\right)=\prod_{i=1}^{n}(a_{i1}a_{i2}\cdots a_{in}),$$

$$\prod_{1\leqslant j<i\leqslant n}a_{ij}=\prod_{i=2}^{n}\prod_{j=1}^{i-1}a_{ij},$$

二、用行列式求非异阵的逆阵

我们可把(1)、(2)中诸式分别合并成矩阵等式如下：

$$\begin{pmatrix} A_{11} & A_{21}\cdots A_{n1} \\ A_{12} & A_{22}\cdots A_{n2} \\ \cdots & \cdots \\ A_{1n} & A_{2n}\cdots A_{nn} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots & \cdots \\ a_{n1} & a_{n2}\cdots a_{nn} \end{pmatrix} = |A|\cdot I_n.$$

$$\begin{pmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots & \cdots \\ a_{n1} & a_{n2}\cdots a_{nn} \end{pmatrix}\begin{pmatrix} A_{11} & A_{21}\cdots A_{n1} \\ A_{12} & A_{22}\cdots A_{n2} \\ \cdots & \cdots \\ A_{1n} & A_{2n}\cdots A_{nn} \end{pmatrix} = |A|\cdot I_n.$$

记

$$\operatorname{adj} A = \begin{pmatrix} A_{11} & A_{21}\cdots A_{n1} \\ A_{12} & A_{22}\cdots A_{n2} \\ \cdots & \cdots \\ A_{1n} & A_{2n}\cdots A_{nn} \end{pmatrix},$$

并把它称为A的**伴随阵**。注意：A_{ij} 在 $\operatorname{adj} A$ 的第(j, i)位置上。由定理 4.1 立刻得出重要等式

$$A(\operatorname{adj} A) = (\operatorname{adj} A)A = |A|\cdot I_n. \tag{10}$$

特别，把它应用于非异阵就得到

定理 4.2 若A为n阶非异阵，则

$$A^{-1} = \frac{1}{|A|}\operatorname{adj} A . \tag{11}$$

这就是用行列式求非异阵的逆阵的一个方法。

例 4 因为

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & -3 \\ 0 & 1 & 1 \end{pmatrix}$$

的行列式 $|A| = 3 \neq 0$，所以A有逆阵。逐个算出代数余子式 A_{ij} 就得

$$A^{-1} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 5 & -2 & -1 \\ -1 & 1 & 2 \\ 1 & -1 & 1 \end{pmatrix}.$$

例 5 设 $A=(a_{ij})$ 是 n 阶正交阵，A_{ij} 是 a_{ij} 的代数余子式，则当 $|A|=1$ 时，$a_{ij}=A_{ij}$；当 $|A|=-1$ 时，$a_{ij}=-A_{ij}$，$i, j=1, 2, \cdots, n$。

【证明】 因为 A 是正交阵，所以 $|A|=\pm 1$。于是

$$A'=A^{-1}=|A|^{-1}\operatorname{adj} A=\pm \operatorname{adj} A.$$

比较上式两端的第(j, i)位置元素即得 $a_{ij}=\pm A_{ij}$，$i, j=1, 2, \cdots, n$，且所取的符号就是$|A|$的符号。 证毕。

主对角元都是 1 的上(下)三角阵称为**单位上(下)三角阵**，它们必是非异阵。

例 6 非异下(上)三角阵的逆阵仍是非异下(上)三角阵。单位下(上)三角阵的逆阵仍是单位下(上)三角阵。

【证明】 今就 A 是下三角阵的情形证明。设

$$A=\begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots\cdots & a_{nn} \end{pmatrix}$$

是任一非异下三角阵，则由定理 3.2 可知 $a_{ii}\neq 0$，$i=1, 2, \cdots, n$。于是

$$A_{11}=a_{22}\cdots a_{nn},$$

$$A_{ii}=a_{11}\cdots a_{i-1,i-1}\cdot a_{i+1,i+1}\cdots a_{nn}\neq 0, \quad 1<i<n,$$

$$A_{nn}=a_{11}\cdots a_{n-1,n-1}.$$

特别，如果所有 $a_{ii}=1$，则所有 $A_{ii}=1$，$i=1, 2, \cdots, n$。

把 A 的第 j 行和第 i 行的元素写出，根据代数余子式的定义，不难发现，当 $i>j$ 时，a_{ij} 的代数余子式 A_{ij} 必定是一个主对角线上出现零的下三角行列式(事实上，A_{ij} 的第(j, j)元素就是 A 的第$(j, j+1)$元素零)，因此，$A_{ij}=0, i>j$，而它是在 A^{-1} 的第(j, i)位置上，因此，A^{-1} 必是下三角阵。

关于上三角阵的结论可类似证明。 证毕。

三、Cramer 法则

本段将解决在§1 中所提出的，用行列式解线性方程组的问

题．先把形如§1中(3)式的n阶线性方程组改写成如下的**矩阵方程**的形状：

$$Ax=b, \tag{12}$$

其中$A=(a_{ij})$为n阶阵，称为(12)的**系数矩阵**，$x=(x_1, x_2, \cdots, x_n)'$和$b=(b_1, b_2, \cdots, b_n)'$为$n$维列向量，分别称为(12)的**未知向量**和**常数向量**．如果(12)有解

$$x_1=c_1, x_2=c_2, \cdots, x_n=c_n,$$

即把它们代入(12)后使其成为恒等式，则称$c=(c, c_2, \cdots, c_n)'$为(12)的一个**解向量**或**解**．若

$$Bx=d \tag{13}$$

是另一个n阶线性方程组．如果(12)的解都是(13)的解，(13)的解都是(12)的解，则称(12)和(13)是**同解方程组**，或称它们是**等价**的．例如，若P是n阶非异阵，则$Ax=b$和$PAx=Pb$是同解方程组．

命题1 若A是非异阵，则(12)必有唯一解$x=A^{-1}b$．

【证明】 把n维列向量$x=A^{-1}b$代入(12)，即知它是(12)的一个解．如果y也是(12)的解：$Ay=b$，则两端各左乘A^{-1}知y必是$y=A^{-1}b$的形状．故$x=A^{-1}b$是(12)的唯一解． 证毕．

定理4.3 (**Cramer法则**)设$A=(\alpha_1, \alpha_2, \cdots, \alpha_n)$是$n$阶阵．如果$|A|\neq 0$，则$n$阶线性方程组$Ax=b$必有唯一解$x=(x_1, x_2, \cdots, x_n)'$，其中第$j$个分量为

$$x_j=\frac{1}{|A|}D_j, \quad j=1, 2, \cdots, n, \tag{14}$$

而

$$D_j=|\alpha_1, \cdots, \alpha_{j-1}, b, \alpha_{j+1}, \cdots, \alpha_n|$$

是用b代替A中第j列α_j所得的n阶行列式．

【证明】 根据定理4.1，D_j按其第j列的展开式为

$$D_j=b_1A_{1j}+b_2A_{2j}+\cdots+b_nA_{nj},$$

但由命题1知

$$x=A^{-1}b=\frac{1}{|A|}(\operatorname{adj} A)b$$

是 $Ax=b$ 的唯一解，比较两端的第 j 个分量就得到

$$x_j=\frac{1}{|A|}(b_1A_{1j}+b_2A_{2j}+\cdots+b_nA_{nj})=\frac{1}{|A|}D_j,$$

$j=1, 2, \cdots, n$。 证毕。

易见，$Ax=b$ 的解的公式(14)就是 §1 中所说的 $n=2, 3$ 时解的公式的推广。

推论 若 $|A|\neq 0$，则 n 阶齐次线性方程组 $Ax=0$ 只有零解。

四、非异阵的分解

作为行列式乘法规则和行列式的展开式的一个应用，我们有

定理 4.4 任何 n 阶非异阵 A 必有分解式

$$A=B\cdot D_n(|A|),$$

这里 B 是有限个 n 阶第三种初等阵的乘积，而

$$D_n(|A|)=[1, \cdots, 1, |A|]$$

为 n 阶对角阵。

【证明】 首先证明，仅用第三种初等行变换可将 A 化为如下形状的上三角阵

$$\begin{pmatrix} 1 & & & * & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & \mu \end{pmatrix} \tag{15}$$

其中 $\mu\neq 0$。设 $A=(a_{ij})_{n\times n}$，因 A 非异，其第 1 列元素不全为零，可不妨设 $a_{11}\neq 0$。因为否则的话，必有某个 $a_{i1}\neq 0(i\neq 1)$，将 A 的第 i 行加到第 1 行上去(这是施行一次第三种初等变换)，所得矩阵的(1, 1)元素必定不为零。由于 $a_{11}\neq 0$，可将第 1 行的 $-\frac{a_{i1}}{a_{11}}$ 倍加到第 i 行上去，从而把第 $(i, 1)$ 元素化为零 $(i=2, \cdots, n)$。如果 $a_{11}\neq 1$，将第 1 行的 $\frac{1-a_{11}}{a_{11}}$ 倍加到第 2 行上，再将所得的第 2 行加到第 1 行上，然后将所得的第 1 行的 $a_{11}-1$ 倍加到第 2 行上，则所得矩阵的(1, 1)元素必为 1，且仍保持所有的$(i, 1)$元素为零

($i=2,3,\cdots\cdots,n$)。至此，仅用第三种初等行变换就把A化成

$$\begin{pmatrix}1 & * \\ 0 & A_1\end{pmatrix}.$$

由行列式展开公式易得$|A|=|A_1|\neq 0$，所以A_1为$n-1$阶非异阵，对A_1可作同样讨论。……如此重复作$n-1$次变换后，A就化成了(15)的形状，且$\mu\neq 0$。

其次，形如(15)的上三角阵，由于$\mu\neq 0$，可将其第n行的$-\frac{a_{in}}{\mu}$倍加到第i行上去，$i=1,2,\cdots,n-1$。因而把第n列中的前$n-1$个元素都化为零。继续对第$n-1$列、第$n-2$列、……、第2列作类似处理(此时代替μ的都是1)可将(15)式化为对角阵

$$D_n(\mu)=[1,\cdots,1,\mu].$$

记所施行的第三种初等行变换的矩阵的乘积为B，则有

$$A=B\cdot D_n(\mu),\ |B|=1.$$

再利用行列式乘法规则即得$\mu=|A|$。　　证毕。

习　　题

1. 试用行列式性质和展开式计算下列行列式

(i) $\begin{vmatrix} 0 & 1 & 2 & -1 & 4 \\ 2 & 0 & 1 & 2 & 1 \\ -1 & 3 & 5 & 1 & 2 \\ 3 & 3 & 1 & 2 & 1 \\ 2 & 1 & 0 & 3 & 5 \end{vmatrix}$；(ii) $\begin{vmatrix} a_0 & -1 & & & \\ a_1 & x & -1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ a_{n-1} & 0 & \cdots & x & -1 \\ a_n & 0 & \cdots & 0 & x \end{vmatrix}$.

2. 用伴随矩阵求下列方阵的逆阵

(i) $\begin{pmatrix}4 & -2 \\ 2 & -3\end{pmatrix}$；(ii) $\begin{pmatrix}1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2\end{pmatrix}$.

3. 设A是n阶复阵，求证

(i) $\mathrm{adj}\,\bar{A}=\overline{\mathrm{adj}\,A}$；

(ii) $\mathrm{adj}(kA)=k^{n-1}\mathrm{adj}\,A$，这里$k$是任一复数；

(iii) $\mathrm{adj}(A')=(\mathrm{adj}\,A)'$；

(iv) 当A是非异阵时，$\mathrm{adj}(A^{-1})=(\mathrm{adj}\,A)^{-1}$。

4. 求证

(i) 若A是对称阵，则 $\mathrm{adj}\,A$ 也是对称阵；

(ii) 若A是偶数阶反对称阵，则 $\mathrm{adj}\,A$ 也是反对称阵。

5. 设 $A=(a_{ij})$是 n 阶对合阵，A_{ij} 是 a_{ij} 的代数余子式，求证

$$A_{ij}=|A|\cdot a_{ji}=\pm a_{ji}\quad (i,\ j=1,\ 2,\ \cdots,\ n).$$

6. 设A是奇数阶非异实方阵，且它的任一元素 a_{ij} 都等于它的代数余子式 A_{ij}。求证$|A|=1$，且A必是正交阵。

7. 设A是 n 阶阵，$n\geqslant 2$。求证

$$|\mathrm{adj}A|=|A|^{n-1}.$$

8. 解下列线性方程组

$$\text{(i)}\ \begin{cases} 3x_1-4x_2-3x_3=4 \\ -x_1+x_2+5x_3=3 \\ 2x_1+3x_2-x_3=-2; \end{cases}$$

$$\text{(ii)}\ \begin{cases} 2x_1-x_2+3x_3+2x_4=6 \\ 3x_1-3x_2+3x_3+2x_4=5 \\ 3x_1-x_2-x_3+2x_4=3 \\ 3x_1-x_2+3x_3-x_4=4. \end{cases}$$

9. 求证如下 n 阶下三角形方程组

$$\begin{cases} x_1=\dfrac{1}{2} \\ \dfrac{1}{2}x_1-x_2=\dfrac{1}{3} \\ \dfrac{1}{3}x_1-\dfrac{1}{2}x_2+x_3=\dfrac{1}{4} \\ \dfrac{1}{4}x_1-\dfrac{1}{3}x_2+\dfrac{1}{2}x_3-x_4=\dfrac{1}{5} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{1}{n}x_1-\dfrac{1}{n-1}x_2+\dfrac{1}{n-2}x_3-\dfrac{1}{n-3}x_4+\cdots+(-1)^{n-2}\dfrac{1}{2}x_{n-1}+(-1)^{n-1}x_n \\ \quad=\dfrac{1}{n+1} \end{cases}$$

的解为$(x_1,\ x_2,\ \cdots,\ x_n)'$，其中 x_k 为如下 k 阶行列式

$$x_k=\begin{vmatrix}\frac{1}{2} & 1 & & & \\ \frac{1}{3} & \frac{1}{2} & 1 & & \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & 1 & \\ \vdots & \vdots & \vdots & & \ddots \\ \frac{1}{k+1} & \frac{1}{k} & \frac{1}{k-1} & \cdots\cdots & \frac{1}{2}\end{vmatrix},\quad k=2,\ \cdots,\ n,$$

而

$$x_1=\frac{1}{2}.$$

10. 求解矩阵方程

$$\begin{pmatrix}1 & 1\\ 3 & 4\end{pmatrix}\begin{pmatrix}x_1 & x_2 & x_3\\ y_1 & y_2 & y_3\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 1\\ 1 & 3 & 4\end{pmatrix}=\begin{pmatrix}-1 & 2 & 1\\ 1 & 3 & -1\end{pmatrix}.$$

11. 如果 n 阶方程组 $Ax=b$ 对任何 n 维列向量 b 都有解，求证：对每个给定的 b，$Ax=b$ 必有唯一解。

12. 求证：行列式为 1 的方阵必可表成有限个第三种初等阵的乘积。

§5 行列式的降阶定理

本节所证明的一些结论在行列式理论中占有重要地位，在本书的不少章节中都要用到它们。

一、分块行列式

把 n 阶阵 A 作如下分块：$A=(A_{ij})_{s\times s}$。今将行列式 $|A|$ 看作以子矩阵 A_{ij} 为元素的所谓 s **阶分块行列式**，记为 $|A|=|A_{ij}|_{s\times s}$。以后常要处理涉及分块行列式的问题。在实用上，常取主对角块 A_{ii} 都是方阵的情形。

命题 1 设 $|A|=|A_{ij}|_{s\times s}$ 是 s 阶分块行列式，则以非零阵 K 左乘（右乘）$|A_{ij}|_{s\times s}$ 的某一行（列）加到另一行（列）上去所得到的新的分块行列式与原分块行列式相等。

【证明】 今就对 $|A_{ij}|_{s\times s}$ 施行行变换来证明。设用 $K\neq 0$ 左乘

分块阵 $(A_{ij})_{s\times s}$ 的第 j 行后加到第 i 行上去，得到的新的分块阵记为 B. 用"八字规则"可知

$$B=T_{ij}(K)(A_{\cdot j})_{s\times s}.$$

再由行列式乘法规则即得

$$|B|=|T_{ij}(K)|\cdot|A_{ij}|_{s\times s}=|A_{ij}|_{s\times s}.$$

对列的变换可仿此证明. 证毕.

例 1 设 A 与 B 是同阶方阵,则

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix}=|A+B|\cdot|A-B|. \tag{1}$$

【证明】 作如下的分块阵初等变换,并应用命题 1 可得

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix}=\begin{vmatrix} A+B & B+A \\ B & A \end{vmatrix}=\begin{vmatrix} A+B & 0 \\ B & A-B \end{vmatrix}$$

（第 2 行乘 (I) 加到第 1 行；第 1 列乘 $(-I)$ 加到第 2 列）

$$=|A+B|\cdot|A-B|.$$

证毕.

用(1)容易算出如下行列式

$$\begin{vmatrix} -a & b & c & d \\ b & -a & d & c \\ c & d & -a & b \\ d & c & b & -a \end{vmatrix}=\begin{vmatrix} c-a & b+d \\ b+d & c-a \end{vmatrix}\cdot\begin{vmatrix} -c-a & b-d \\ b-d & -c-a \end{vmatrix}$$

$$=\{(a-c)^2-(b+d)^2\}\cdot\{(a+c)^2-(b-d)^2\}$$

$$=(a+b-c+d)(a-b-c-d)(a+b+c-d)(a-b+c+d).$$

例 2 设 B 与 C 是同阶实方阵,则

$$\begin{vmatrix} B & C \\ -C & B \end{vmatrix}=|B+\sqrt{-1}C|\cdot|B-\sqrt{-1}C|. \tag{2}$$

【证明】

$$\begin{vmatrix} B & C \\ -C & B \end{vmatrix}=\begin{vmatrix} B+\sqrt{-1}C & C \\ \sqrt{-1}(B+\sqrt{-1}C) & B \end{vmatrix}$$

（第 2 列乘 $(\sqrt{-1}I)$ 加到第 1 列；第 1 行乘 $(-\sqrt{-1}I)$ 加到第 2 行）

$$=\begin{vmatrix} B+\sqrt{-1}C & C \\ 0 & B-\sqrt{-1}C \end{vmatrix}=|B+\sqrt{-1}C|\cdot|B-\sqrt{-1}C|.$$

证毕。

同样，用(2)容易算出如下行列式

$$\begin{vmatrix} a & b & c & d \\ b & a & -d & c \\ -c & -d & a & b \\ d & -c & b & a \end{vmatrix} = (a^2-b^2+c^2-d^2)^2,$$

(2) 式还可用来把 n 阶复线性方程组 $Ax=b$ 化为与它同解的 $2n$ 阶实线性方程组。令

$$x=y+\sqrt{-1}z,\quad A=B+\sqrt{-1}C,\quad b=\beta+\sqrt{-1}\delta,$$

其中 y, z 是未知的实 n 维列向量，B 与 C 分别是 A 的实部(矩阵)与虚部(矩阵)①，β 与 δ 分别是 b 的实部(向量)与虚部(向量)，于是由

$$(B+\sqrt{-1}C)(y+\sqrt{-1}z)=\beta+\sqrt{-1}\delta,$$

可得

$$\begin{cases} By-Cz=\beta \\ Cy+Bz=\delta. \end{cases}$$

也就是说，$2n$ 阶实方程组

$$\begin{pmatrix} B & -C \\ C & B \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}=\begin{pmatrix} \beta \\ \delta \end{pmatrix}$$

与 $Ax=b$ 同解。由(2)式可知，此方程的系数行列式恰好等于

$$|\overline{A}|\cdot|A|=\overline{|A|}\cdot|A|=|\det A|^2.$$

二、行列式的两个降阶定理

定理 5.1 （**第一降阶定理**）② 设 $M=\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 是方阵，其中 A 是非异阵，则

---

① $B=\frac{1}{2}(A+\overline{A})$，$C=\frac{1}{2\sqrt{-1}}(A-\overline{A})$.

② 一些矩阵论专著及近期的文献中常称这个定理为 Schur 定理，但也有称为加株(カおこ)定理的。为明确其实际含义，我们改称为第一降阶定理。

$$|M| = \begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \cdot |D - CA^{-1}B| = |A| \cdot |M/A|, \qquad (3)$$

其中 $M/A = D - CA^{-1}B$，称为 **Schur 补**。

【证明】 这是下式的直接推论：

$$\begin{matrix} (-CA^{-1}) \\ \longrightarrow \end{matrix} \begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} A & B \\ 0 & D - CA^{-1}B \end{vmatrix}.$$

证毕。

推论 1 设 A 是 n 阶非异阵，D 是 m 阶阵，B 与 C 分别是 $n \times m$ 阵与 $m \times n$ 阵，则有如下**升阶公式**：

$$|D - CA^{-1}B| = \frac{1}{|A|} \begin{vmatrix} A & B \\ C & D \end{vmatrix}.$$

推论 2 设 A, B, C, D 是同阶方阵，如果 A 是非异阵，且 $AC = CA$，则

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - CB|. \qquad (4)$$

【证明】 由第一降阶定理，并应用行列式乘法规则得

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| \cdot |D - CA^{-1}B| = |AD - ACA^{-1}B|.$$

由假设 $AC = CA$，即可化为(4)式。 证毕。

实际上，推论 2 中 A 的非异性条件是可以去掉的，即使 A 是奇异阵，(4)式仍然成立，这将在第四章中证明。

第一降阶定理的另一种形式是下面的

定理 5.1′ 设 $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 是方阵，其中 D 是非异阵，则

$$|M| = |D| \cdot |A - BD^{-1}C|.$$

由于证法与定理 5.1 的类似，故从略。

若把定理 5.1′ 与升阶公式合并运用，就得到另一种实用的降阶定理。

定理 5.2 （**第二降阶定理**）设 A 与 D 分别是 n 阶与 m 阶非异阵，B 与 C 分别是 $n \times m$ 阵与 $m \times n$ 阵，则

$$|D-CA^{-1}B|=\frac{|D|}{|A|}\cdot|A-BD^{-1}C|. \tag{5}$$

行列式的两个降阶定理的证明虽然是一目了然的，然而它们却有着极其广泛的应用。下面举五个例子说明之。

例 3 计算行列式

$$D=\begin{vmatrix}2 & 1 & 0 & 1 & -1\\ 3 & 2 & 1 & 0 & 2\\ 1 & 1 & 2 & -1 & 5\\ 0 & -1 & 6 & 1 & 4\\ 4 & 5 & 8 & 7 & -1\end{vmatrix}.$$

【解】 由第一降阶定理得

$$D=\left|\begin{pmatrix}2 & -1 & 5\\ 6 & 1 & 4\\ 8 & 7 & -1\end{pmatrix}-\begin{pmatrix}1 & 1\\ 0 & -1\\ 4 & 5\end{pmatrix}\begin{pmatrix}2 & 1\\ 3 & 2\end{pmatrix}^{-1}\begin{pmatrix}0 & 1 & -1\\ 1 & 0 & 2\end{pmatrix}\right|$$

$$=\begin{vmatrix}1 & 0 & 2\\ 8 & -2 & 11\\ 2 & 14 & -20\end{vmatrix}=\left|\begin{pmatrix}-2 & 11\\ 14 & -20\end{pmatrix}-\begin{pmatrix}8\\ 2\end{pmatrix}(0\quad 2)\right|$$

$$=\begin{vmatrix}-2 & -5\\ 14 & -24\end{vmatrix}=118.$$

例 4 求证：实镜象阵的行列式 $|I_n-2uu'|=-1$，这里设 $u'u=1$，u 是 n 维列向量。

【证明】 由第二降阶定理即得

$$|I_n-2uu'|=|I_n-(2u)\cdot 1^{-1}\cdot u'|=\frac{|I_n|}{1}(1-2u'u)=-1.$$

证毕。

这个证明可以说是求实镜象阵的行列式的最简捷的方法了。

为了运用时方便，常把第二降阶定理的(5)式写成

$$|D+CA^{-1}B|=\frac{|D|}{|A|}\cdot|A+BD^{-1}C|. \tag{6}$$

例 5 设 A 是 n 阶非异阵，α 和 β 是两个 n 维列向量，则由(6)式可得

$$|A+\alpha\beta'|=|A|\cdot(1+\beta'A^{-1}\alpha).$$

所以，$A+\alpha\beta'$ 是非异阵当且仅当 $1+\beta'A^{-1}\alpha\neq 0$。

例 6 设 $\prod_{i=1}^{n} a_i\neq 0$。计算行列式

$$D=\begin{vmatrix} 0 & a_1+a_2 & \cdots\cdots & a_1+a_n \\ a_2+a_1 & 0 & \cdots\cdots & a_2+a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n+a_1 & a_n+a_2 & \cdots\cdots & 0 \end{vmatrix},\quad n\geqslant 2.$$

【解】 先把 D 改写成

$$D=\left|\begin{pmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n \end{pmatrix}+\begin{pmatrix} a_1+a_1 & a_1+a_2 & \cdots & a_1+a_n \\ a_2+a_1 & a_2+a_2 & \cdots & a_2+a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_n+a_1 & a_n+a_2 & \cdots & a_n+a_n \end{pmatrix}\right|$$

$$=\left|\begin{pmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n \end{pmatrix}\right.$$

$$\left.+\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 1 & \cdots\cdots & 1 \\ a_1 & a_2 & \cdots\cdots & a_n \end{pmatrix}\right|,$$

故由(6)式即得

$$D=\frac{\begin{vmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n \end{vmatrix}}{\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}}$$

$$\times\left|\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}\begin{pmatrix} -2a_1 & & & \\ & -2a_2 & & \\ & & \ddots & \\ & & & -2a_n \end{pmatrix}^{-1}\begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}\right|$$

$$=(-2)^n\prod_{i=1}^n a_i\cdot\begin{vmatrix}1-\dfrac{n}{2} & -\dfrac{1}{2}\sum\limits_{i=1}^n\dfrac{1}{a_i}\\ -\dfrac{1}{2}\sum\limits_{j=1}^n a_j & 1-\dfrac{n}{2}\end{vmatrix}$$

$$=(-2)^{n-2}\prod_{i=1}^n a_i\left((n-2)^2-\sum_{i,j=1}^n\frac{a_j}{a_i}\right).$$

例 7 设 $A=(a_{ij})_{n\times n}$ 是 n 阶对称阵，令

$$\Delta_i=\left|A\begin{pmatrix}1 & 2\cdots i\\ 1 & 2\cdots i\end{pmatrix}\right|,\ i=1,2,\cdots,n.$$

对任意一个 i: $1<i<n$，如果 $\Delta_i=0$ 而 $\Delta_{i-1}\cdot\Delta_{i+1}\neq 0$，求证：$\Delta_{i-1}\cdot\Delta_{i+1}<0$。

【证明】 记 A 的第 $i-1$ 个顺序主子阵为

$$A_{i-1}=A\begin{pmatrix}1 & 2\cdots i-1\\ 1 & 2\cdots i-1\end{pmatrix}.$$

将 Δ_i 写成分块行列式

$$\Delta_i=\begin{vmatrix}A_{i-1} & \alpha\\ \alpha' & a_{ii}\end{vmatrix}. \tag{7}$$

由假设 $\Delta_{i-1}\neq 0$，故由第一降阶定理，(7)式可化为

$$\Delta_i=\Delta_{i-1}\cdot(a_{ii}-\alpha' A_{i-1}^{-1}\alpha).$$

再由假设 $\Delta_i=0$，故得

$$a_{ii}-\alpha' A_{i-1}^{-1}\alpha=0. \tag{8}$$

同理，由第一降阶定理可知

$$\Delta_{i+1}=\begin{vmatrix}A_{i-1} & \alpha & \beta\\ \alpha' & a_{ii} & a_{i,i+1}\\ \beta' & a_{i,i+1} & a_{i+1,i+1}\end{vmatrix}$$

$$=\Delta_{i-1}\left|\begin{pmatrix}a_{ii} & a_{i,i+1}\\ a_{i,i+1} & a_{i+1,i+1}\end{pmatrix}-\begin{pmatrix}\alpha'\\ \beta'\end{pmatrix}A_{i-1}^{-1}(\alpha\ \ \beta)\right|$$

$$=\Delta_{i-1}\begin{vmatrix}a_{ii}-\alpha' A_{i-1}^{-1}\alpha & a_{i,i+1}-\alpha' A_{i-1}^{-1}\beta\\ a_{i,i+1}-\beta' A_{i-1}^{-1}\alpha & a_{i+1,i+1}-\beta' A_{i-1}^{-1}\beta\end{vmatrix}. \tag{9}$$

因为 A 是对称阵，故 A_{i-1} 和 A_{i-1}^{-1} 也是对称阵。记

$$g = a_{i,i+1} - \beta' A_{i-1}^{-1}\alpha,$$

于是

$$g = g' = a_{i,i+1} - \alpha' A_{i-1}^{-1}\beta.$$

由(8)式知(9)式可化为

$$\Delta_{i+1} = \Delta_{i-1}\begin{vmatrix} 0 & g \\ g & * \end{vmatrix} = \Delta_{i-1}(-g^2).$$

但已知 $\Delta_{i-1}\cdot\Delta_{i+1}\neq 0$，所以必有 $\Delta_{i-1}\cdot\Delta_{i+1}<0$。　　　证毕。

例 4～例 6 表明，运用第二降阶定理，可把计算 n 阶行列式的问题归结为计算一阶或二阶行列式。特别是例 6 给我们以启示：若能把 n 阶行列式 $|A|$ 对应的矩阵 A 分解为

$$A = D + M, \tag{10}$$

这里 D 是一个比较容易计算行列式和逆阵的非异阵（例如对角阵等），而

$$M = HL, \tag{11}$$

其中 H 是 $n\times r$ 阵，L 是 $r\times n$ 阵，$r<n$，则由(6)式可知

$$|A| = |D+HL| = |D|\cdot|I_r + LD^{-1}H|. \tag{12}$$

于是就把计算 n 阶行列式化为计算 r 阶行列式。那么，怎样找分解式(10)与(11)，且使 r 尽可能小呢？这是个有趣的问题。关于分解式(11)的问题，在第三章中将涉及到。

习　题

1. 设 A 和 C 分别是 m 阶和 n 阶非异阵，求证 $m+n$ 阶阵 $\begin{pmatrix} 0 & A \\ C & B \end{pmatrix}$ 必是非异阵，并求出其逆阵。

2. (Szaraki-Wazewski) 设 A，B 是同阶方阵，求证

$$\begin{vmatrix} \lambda I - A & B \\ -B & \lambda I - A \end{vmatrix} = |\lambda I - (A+\sqrt{-1}B)|\cdot|\lambda I - (A-\sqrt{-1}B)|.$$

(这是例 2 的(2)式的推广.)

3. 设 $AB=BA$，求证 $\begin{vmatrix} A & -B \\ B & A \end{vmatrix} = |A^2+B^2|$。

4. 设 A 与 B 是同阶方阵，计算行列式

(i) $\begin{vmatrix} A & B \\ B & -A \end{vmatrix}$； (ii) $D=\begin{vmatrix} A & JBJ \\ B & JAJ \end{vmatrix}$，这里 $J=\begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}$。

(称D为中心对称行列式。)

5. 设 A, B, C, D 都是同阶方阵，求证

$$\begin{vmatrix} A & B & C & D \\ B & A & D & C \\ C & D & A & B \\ D & C & B & A \end{vmatrix}$$

$$=|A+B+C+D|\cdot|A+B-C-D|\cdot|A-B+C-D|$$
$$\cdot|A-B-C+D|.$$

6. 用第一降阶定理计算下面的行列式

(i) $\begin{vmatrix} 1 & -2 & 1 & 1 \\ -2 & 4 & 3 & -3 \\ 5 & 1 & 0 & 7 \\ 3 & 1 & 4 & 2 \end{vmatrix}$； (ii) $\begin{vmatrix} 1 & 2 & -1 & 3 & -5 \\ -2 & -4 & -2 & 1 & 2 \\ -1 & 7 & -3 & 4 & 0 \\ 0 & 5 & 6 & 1 & 1 \\ 6 & -1 & 0 & 3 & 7 \end{vmatrix}$；

(iii) $\begin{vmatrix} 0 & 1 & 2 & \cdots\cdots & n \\ 1 & 1 & & & \\ 2 & & 2 & & \\ \vdots & & & \ddots & \\ n & & & & n \end{vmatrix}$；

(iv) $\begin{vmatrix} a_{11} & a_{12}\cdots a_{1n} & -y_1 \\ \cdots\cdots & \cdots\cdots & \cdots \\ a_{n1} & a_{n2}\cdots a_{nn} & -y_n \\ x_1 & x_2\cdots\cdots x_n & 0 \end{vmatrix}$，这里假设，$\det(a_{ij})_{n\times n}\neq 0$。

7. 计算下面的行列式

(i) $\begin{vmatrix} 1 & b & b & \cdots\cdots & b \\ c & \lambda & d & \cdots\cdots & d \\ c & d & \lambda & \cdots\cdots & d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c & d & d & \cdots\cdots & \lambda \end{vmatrix}$； (ii) $\begin{vmatrix} 1 & & & 1 & & \\ & \ddots & & & \ddots & \\ & & 1 & & & 1 \\ -\frac{1}{2} & & & \frac{1}{2} & & \\ & \ddots & & & \ddots & \\ & & -\frac{1}{2} & & & \frac{1}{2} \end{vmatrix}$。

8. 计算下面的行列式

(i) $\begin{vmatrix} A+aI_n & A+bI_n \\ A+bI_n & A+aI_n \end{vmatrix}$, $b\neq 0$;

(ii) $\begin{vmatrix} aI_n & bJ_n \\ bJ_n & aI_n \end{vmatrix}$, $J_n=\begin{pmatrix} & & & 1 \\ & & 1 & \\ & ⋰ & & \\ 1 & & & \end{pmatrix}_{n\times n}$;

(iii) $\begin{vmatrix} I_n & J_n \\ J_n & -I_n \end{vmatrix}$.

9. 设S是$2n$阶非异反对称阵，α是$2n$维实列向量，求证$\begin{pmatrix} S & \alpha \\ \alpha' & k \end{pmatrix}$为非异阵的充分必要条件是$k\neq 0$。

10. 用第二降阶定理计算下面的行列式。

(i) $\begin{vmatrix} 0 & 2 & 3 & \cdots\cdots & n \\ 1 & 0 & 3 & \cdots\cdots & n \\ 1 & 2 & 0 & \cdots\cdots & n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 2 & 3 & \cdots\cdots & 0 \end{vmatrix}$; (ii) $\begin{vmatrix} 2 & 1 & 1 & \cdots\cdots & 1 \\ 1 & 3 & 1 & \cdots\cdots & 1 \\ 1 & 1 & 4 & \cdots\cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots\cdots & n+1 \end{vmatrix}$;

(iii) $\begin{vmatrix} \lambda_1 & a_2 & a_3 & \cdots\cdots & a_n \\ a_1 & \lambda_2 & a_3 & \cdots\cdots & a_n \\ a_1 & a_2 & \lambda_3 & \cdots\cdots & a_n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_1 & a_2 & a_3 & \cdots\cdots & \lambda_n \end{vmatrix}$;

(iv) $\begin{vmatrix} a_1^2 & a_1a_2+1 & \cdots\cdots & a_1a_n+1 \\ a_2a_1+1 & a_2^2 & \cdots\cdots & a_2a_n+1 \\ \cdots & \cdots & \cdots & \cdots \\ a_na_1+1 & a_na_2+1 & \cdots\cdots & a_n^2 \end{vmatrix}$.

11. 设A是$m\times n$阵，B是$n\times m$阵，$m\geqslant n$，$\lambda\neq 0$，求证

$$|\lambda I_m-AB|=\lambda^{m-n}|\lambda I_n-BA|.$$

§6 Laplace 定理、两个方阵之和的行列式

主要出于在理论推导中的需要，本段将把 §4 中行列式按某一

列(或行)的展开式加以推广。

设 $A=(a_{ij})_{n\times n}$ 是 n 阶阵，与一阶子式 a_{ij} 在 $|A|$ 中的代数余子式 A_{ij} 的概念相仿，我们要引进 k 阶子式的代数余子式的概念，$1\leqslant k\leqslant n$。

定义 设 $A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}$ 是 $A=(a_{ij})_{n\times n}$ 的 k 阶子矩阵，则称它的行列式 $\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$ 为 $|A|$ 的(或 A 的)k **阶子式**。特别，称 $\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ i_1 & i_2\cdots i_k \end{pmatrix}\right|$ 为 k **阶主子式**，$\left|A\begin{pmatrix} 1 & 2\cdots k \\ 1 & 2\cdots k \end{pmatrix}\right|$ 为 k **阶顺序主子式**，这里 $1\leqslant k\leqslant n$，$1\leqslant i_1<i_2<\cdots<i_k\leqslant n$，$1\leqslant j_1<j_2<\cdots<j_k\leqslant n$。

定义 设 $\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$ 是 n 阶行列式 $|A|$ 的 k 阶子式，$1\leqslant k\leqslant n-1$。在 $|A|$ 中划掉第 $i_1, i_2, \cdots, i_k$ 行和第 $j_1, j_2, \cdots, j_k$ 列的元素，剩下的元素按原来相对位置所排成的一个 $n-k$ 阶行列式，记为 $\left|A_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$，并称它为 $\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$ 在 $|A|$ 中的**余子式**，称 $n-k$ 阶行列式

$$\left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|=(-1)^{s+t}\left|A_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$$

为 $\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$ 在 $|A|$ 中的**代数余子式**，这里 $s=\sum\limits_{l=1}^{k} i_l$，$t=\sum\limits_{l=1}^{k} j_l$。

由定义易知，$\left|A_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$ 在 $|A|$ 中的余子式就是 $\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$，因此，这两者互为余子式，任一子式的余子式与代数余子式最多差一个负号。

例 1 在行列式 $|A|=|a_{ij}|_{4\times 4}$ 中，2 阶子式

$$\left|A\begin{pmatrix}2 & 4\\ 1 & 2\end{pmatrix}\right|=\begin{vmatrix}a_{21} & a_{22}\\ a_{41} & a_{42}\end{vmatrix}$$

的余子式是

$$\left|A_c\begin{pmatrix}2 & 4\\ 1 & 2\end{pmatrix}\right|=\left|A\begin{pmatrix}1 & 3\\ 3 & 4\end{pmatrix}\right|=\begin{vmatrix}a_{13} & a_{14}\\ a_{33} & a_{34}\end{vmatrix},$$

而代数余子式是

$$\left|\hat{A}_c\begin{pmatrix}2 & 4\\ 1 & 2\end{pmatrix}\right|=(-1)^{(2+4)+(1+2)}\left|A_c\begin{pmatrix}2 & 4\\ 1 & 2\end{pmatrix}\right|=-\begin{vmatrix}a_{13} & a_{14}\\ a_{33} & a_{34}\end{vmatrix}.$$

引理 $\left|A\begin{pmatrix}i_1 & i_2\cdots i_k\\ j_1 & j_2\cdots j_k\end{pmatrix}\right|\cdot\left|\hat{A}_c\begin{pmatrix}i_1 & i_2\cdots i_k\\ j_1 & j_2\cdots j_k\end{pmatrix}\right|$ 的展开式中的任一项必是$|A|$的展开式中的某一项，而且这两者的符号也相同。

【证明】 先对 $i_1=1,\ i_2=2,\ \cdots,\ i_k=k$; $j_1=1,\ j_2=2,\ \cdots,\ j_k=k$ 这一特殊情形证明之。此时

$$|A|=\begin{vmatrix}A\begin{pmatrix}1 & 2\cdots k\\ 1 & 2\cdots k\end{pmatrix} & * \\ * & A_c\begin{pmatrix}1 & 2\cdots k\\ 1 & 2\cdots k\end{pmatrix}\end{vmatrix}.$$

因为

$$\left|A\begin{pmatrix}1 & 2\cdots k\\ 1 & 2\cdots k\end{pmatrix}\right|=\begin{vmatrix}a_{11} & \cdots & a_{1k}\\ \cdots & \cdots & \cdots\\ a_{k1} & \cdots & a_{kk}\end{vmatrix}$$

的任一项是

$$(-1)^{N(j_1j_2\cdots j_k)}a_{1j_1}a_{2j_2}\cdots a_{kj_k},$$

其中 $j_1j_2\cdots j_k$ 是 $1,2,\cdots,k$ 的某一排列，而

$$\left|A_c\begin{pmatrix}1 & 2\cdots k\\ 1 & 2\cdots k\end{pmatrix}\right|=\left|A\begin{pmatrix}k+1 & k+2\cdots n\\ k+1 & k+2\cdots n\end{pmatrix}\right|$$

$$=\begin{vmatrix}a_{k+1,k+1} & \cdots\cdots & a_{k+1,n}\\ \cdots & \cdots & \cdots\\ a_{n,k+1} & \cdots\cdots & a_{nn}\end{vmatrix}$$

的任一项是

$$(-1)^{N(j_{k+1}j_{k+2}\cdots j_n)}a_{k+1,j_{k+1}}a_{k+2,j_{k+2}}\cdots a_{nj_n},$$

其中 $j_{k+1}j_{k+2}\cdots j_n$ 是$(k+1)$, $(k+2)$, $\cdots$, n 的某一排列。由于

$$\left|\hat{A}_\sigma\begin{pmatrix}1 & 2\cdots k\\1 & 2\cdots k\end{pmatrix}\right| = \left|A_\sigma\begin{pmatrix}1 & 2\cdots k\\1 & 2\cdots k\end{pmatrix}\right|$$

所以

$$\left|A\begin{pmatrix}1 & 2\cdots k\\1 & 2\cdots k\end{pmatrix}\right|\cdot\left|\hat{A}_\sigma\begin{pmatrix}1 & 2\cdots k\\1 & 2\cdots k\end{pmatrix}\right|$$

$$=\left|A\begin{pmatrix}1 & 2\cdots k\\1 & 2\cdots k\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}k+1 & k+2\cdots n\\k+1 & k+2\cdots n\end{pmatrix}\right|$$

的任一项恰是

$$\delta a_{1j_1}a_{2j_2}\cdots a_{kj_k}\cdot a_{k+1,j_{k+1}}\cdots\cdots a_{nj_n},$$

而其符号

$$\delta=(-1)^{N(j_1j_2\cdots j_k)+N(j_{k+1}j_{k+2}\cdots j_n)}$$
$$=(-1)^{N(j_1\cdots j_kj_{k+1}\cdots j_n)}.$$

所以它确是$|A|$的展开式中的某一项，且符号相同。

其次，对一般的

$$1\leqslant i_1<i_2<\cdots<i_k\leqslant n;\ 1\leqslant j_1<j_2<\cdots<j_k\leqslant n,$$

可按如下办法处理，经过 i_1-1 次相邻两行的对调，可把$|A|$的第 i_1 行调成第 1 行。同理，经过 i_2-2, $\cdots$, i_k-k 次相邻两行的对调，把原来的第 i_2, $\cdots$, i_k 行调成第 2, $\cdots$, k 行，于是经过

$$(i_1-1)+(i_2-2)+\cdots+(i_k-k)=\sum_{l=1}^{k}i_l-\sum_{l=1}^{k}l$$

次相邻行的对调，就把$|A|$的第 i_1, i_2, $\cdots$, i_k 行调到前 k 个行的位置上。同理，经过

$$(j_1-1)+(j_2-2)+\cdots+(j_k-k)=\sum_{l=1}^{k}j_l-\sum_{l=1}^{k}l$$

次相邻列的对调，就把$|A|$的第 j_1, j_2, $\cdots$, j_k 列调到前 k 个列的位置上。总起来说，对$|A|$经过

$$\sum_{l=1}^{k}i_l+\sum_{l=1}^{k}j_l-2\sum_{l=1}^{k}l$$

次相邻行或列的对调后，得到一个新的行列式

$$D=\begin{vmatrix} A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix} & * \\ * & B \end{vmatrix},$$

而有

$$D=(-1)^{s+t}|A|,$$

这里 $s=\sum_{l=1}^{k} i_l$, $t=\sum_{l=1}^{k} j_l$. D 中的子式 $|B|$ 表示

$$\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$$

在 D 中的代数余子式. 但由上述行列对调方法显然可知

$$|B|=\left|A_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$$

$$\left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|=(-1)^{s+t}|B|.$$

这就归之为刚才讨论过的第一种情形,所以

$$\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|\cdot|B|$$

的任一项不仅是 D 的展开式中的某一项,且两者符号相同. 于是

$$\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|\cdot\left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|$$

$$=(-1)^{s+t}\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|\cdot|B|$$

的任一项不仅恰是 $(-1)^{s+t}D=|A|$ 的某一项,而且两者符号一致,故引理的结论正确.

定理 6.1(Laplace) 设 A 是 n 阶阵, 在 $|A|$ 中任意取定 k 个列,设其列号合于 $1\leqslant j_1<j_2<\cdots<j_k\leqslant n$, 则

$$|A|=\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|\cdot\left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|. \quad (1)$$

类似地,在 $|A|$ 中任意取定 k 个行, 设其行号合于 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$, 则

$$|A| = \sum_{1\leqslant j_1<j_2<\cdots<j_k\leqslant n} \left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right|. \tag{2}$$

这就是说，任何 n 阶行列式 $|A|$ 等于它的位于同 k 个行(列)的 C_n^k 个 k 阶子式乘以相应的代数余子式之和。

【证明】 以证明(1)式为例，利用 $|A'| = |A|$，即得(2)式。

记(1)的右端为 S。由引理，S 中的任意一个乘积

$$\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right|$$

的展开式中的任一项都是 $|A|$ 的展开式中的项，S 中的任意两个乘积项

$$\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right|$$

与

$$\left|A\begin{pmatrix} l_1 & l_2\cdots l_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} l_1 & l_2\cdots l_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right|$$

的各自的展开式中，由于至少有某个 $1\leqslant p\leqslant k$ 使 $i_p \neq l_p$，所以它们对 $|A|$ 的展开式所提供的项是没有重复的，这样，S 中的总项数为

$$C_n^k \cdot k! \cdot (n-k)! = n!,$$

而这恰好是 $|A|$ 的项数，且相应项的符号都一致，因此，$|A| = S$，即(1)式成立。 证毕。

称(1)式为 $|A|$ 按它的第 $j_1, j_2, \cdots, j_k$ 列的展开式。称(2)式为 $|A|$ 按它的第 $i_1, i_2, \cdots, i_k$ 行的展开式。

例 2 4 阶行列式 $|A|$ 按它的第 2 和第 4 两列的展开式是

$$\begin{aligned}|A| = &\left|A\begin{pmatrix} 1 & 2 \\ 2 & 4\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} 1 & 2 \\ 2 & 4\end{pmatrix}\right| + \left|A\begin{pmatrix} 1 & 3 \\ 2 & 4\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} 1 & 3 \\ 2 & 4\end{pmatrix}\right| \\ &+ \left|A\begin{pmatrix} 1 & 4 \\ 2 & 4\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} 1 & 4 \\ 2 & 4\end{pmatrix}\right| + \left|A\begin{pmatrix} 2 & 3 \\ 2 & 4\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} 2 & 3 \\ 2 & 4\end{pmatrix}\right| \\ &+ \left|A\begin{pmatrix} 2 & 4 \\ 2 & 4\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} 2 & 4 \\ 2 & 4\end{pmatrix}\right| + \left|A\begin{pmatrix} 3 & 4 \\ 2 & 4\end{pmatrix}\right| \cdot \left|\hat{A}_c\begin{pmatrix} 3 & 4 \\ 2 & 4\end{pmatrix}\right|\end{aligned}$$

$$= -\left|A\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}3 & 4\\ 1 & 3\end{pmatrix}\right| + \left|A\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}2 & 4\\ 1 & 3\end{pmatrix}\right|$$

$$-\left|A\begin{pmatrix}1 & 4\\ 2 & 4\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}2 & 3\\ 1 & 3\end{pmatrix}\right| - \left|A\begin{pmatrix}2 & 3\\ 2 & 4\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}1 & 4\\ 1 & 3\end{pmatrix}\right|$$

$$+\left|A\begin{pmatrix}2 & 4\\ 2 & 4\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}1 & 3\\ 1 & 3\end{pmatrix}\right| - \left|A\begin{pmatrix}3 & 4\\ 2 & 4\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}1 & 2\\ 1 & 3\end{pmatrix}\right|.$$

作为 Laplace 定理的一个应用，我们介绍关于两个方阵之和的行列式计算公式，其中一个矩阵是纯量矩阵。

定理 6.2 设 A 是任一 n 阶阵，则

$$|\lambda I_n + A| = \lambda^n + \left(\sum_{i_1=1}^{n}\left|A\begin{pmatrix}i_1\\ i_1\end{pmatrix}\right|\right)\lambda^{n-1}$$

$$+\left(\sum_{1\leq i_1<i_2\leq n}\left|A\begin{pmatrix}i_1 & i_2\\ i_1 & i_2\end{pmatrix}\right|\right)\lambda^{n-2} + \cdots\cdots$$

$$+\left(\sum_{1\leq i_1<i_2<\cdots<i_{n-1}\leq n}\left|A\begin{pmatrix}i_1 & i_2\cdots i_{n-1}\\ i_1 & i_2\cdots i_{n-1}\end{pmatrix}\right|\right)\lambda$$

$$+|A|, \tag{3}$$

即 λ^{n-k} 的系数是 $|A|$ 的所有 k 阶主子式之和。

【证】 写 $\lambda I_n = (\lambda e_1, \lambda e_2, \cdots, \lambda e_n)$，$A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$。连续运用行列式的性质 3 可得

$$|\lambda I_n + A| = |\lambda e_1 + \alpha_1, \lambda e_2 + \alpha_2, \cdots, \lambda e_n + \alpha_n|$$

$$= |\lambda e_1, \lambda e_2, \cdots, \lambda e_n| + \sum_{j=1}^{n}|\lambda e_1 \cdots \lambda e_{j-1}, \alpha_j, \lambda e_{j+1}, \cdots, \lambda e_n|$$

$$+ \sum_{1\leq j_1<j_2\leq n}|\lambda e_1, \cdots, \lambda e_{j_1-1}, \alpha_{j_1}, \lambda e_{j_1+1}, \cdots, \lambda e_{j_2},$$

$$\lambda e_{j_2-1}, \alpha_{j_2}, \lambda e_{j_2+1}, \cdots, \lambda e_n| + \cdots\cdots$$

$$+ \sum_{1\leq j_1<\cdots<j_{n-1}\leq n}|\alpha_{j_1}, \cdots, \lambda e_{j_n}, \cdots, \alpha_{j_{n-1}}|$$

$$+ |\alpha_1\alpha_2\cdots\alpha_n|.$$

应用 Laplace 定理，把上式右边的 $n-1$ 个和号中的行列式分别按其第 j 列，第 j_1、j_2 列，…，第 j_1、j_2、…、j_{n-1} 列展开即得(3)式。

证毕。

习　　题

1. 用 Laplace 定理证明

$$\begin{vmatrix} a_1 & & & & & b_1 \\ & \ddots & & & \iddots & \\ & & a_n & b_n & & \\ & & b_{n+1} & a_{n+1} & & \\ & \iddots & & & \ddots & \\ b_{2n} & & & & & a_{2n} \end{vmatrix} = \prod_{i=1}^{n} (a_i a_{2n-i+1} - b_i b_{2n-i+1}).$$

2. 计算下面的 n 阶行列式

(i) $\begin{vmatrix} \lambda - a_1 b_1 & -a_1 b_2 \cdots\cdots & -a_1 b_n \\ -a_2 b_1 & \lambda - a_2 b_2 \cdots\cdots & -a_2 b_n \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ -a_n b_1 & -a_n b_2 \cdots\cdots & \lambda - a_n b_n \end{vmatrix}$;

(ii) $\begin{vmatrix} \lambda_1 & a_1 b_2 \cdots\cdots & a_1 b_n \\ a_2 b_1 & \lambda_2 \cdots\cdots & a_2 b_n \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_n b_1 & a_n b_2 \cdots\cdots & \lambda_n \end{vmatrix}$;

(iii) $\begin{vmatrix} \lambda + a_1 & a_2 \cdots\cdots & a_n \\ a_1 & \lambda + a_2 \cdots\cdots & a_n \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_1 & a_2 \cdots\cdots & \lambda + a_n \end{vmatrix}$;

(iv) $\begin{vmatrix} \lambda_1 & a_2 + 1 \cdots\cdots & a_n + 1 \\ a_1 + 1 & \lambda_2 \cdots\cdots & a_n + 1 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_1 + 1 & a_2 + 1 \cdots\cdots & \lambda_n \end{vmatrix}$.

3. 设 $a_1, a_2, \cdots, a_n$ 是 n 个不全为零的实数，求证

$$\begin{vmatrix} 1 + a_1^2 & a_1 a_2 \cdots\cdots & a_1 a_n \\ a_2 a_1 & 1 + a_2^2 \cdots\cdots & a_2 a_n \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_n a_1 & a_n a_2 \cdots\cdots & 1 + a_n^2 \end{vmatrix} > 1.$$

4. 求证

$$\begin{vmatrix} \lambda_1+a_{11} & \lambda_2+a_{12}\cdots\cdots\lambda_n+a_{1n} \\ \lambda_1+a_{21} & \lambda_2+a_{22}\cdots\cdots\lambda_n+a_{2n} \\ \cdots\cdots & \cdots\cdots \\ \lambda_1+a_{n1} & \lambda_2+a_{n2}\cdots\cdots\lambda_n+a_{nn} \end{vmatrix} = \sum_{i,j=1}^{n} \lambda_j A_{ij} + |A|.$$

这里$|A|=|a_{ij}|_{n\times n}$，A_{ij}是a_{ij}的代数余子式。

5. 由上题以及行列式的性质证明

$$\begin{vmatrix} a_{11}-a_{12} & a_{12}-a_{13}\cdots\cdots a_{1,n-1}-a_{1n} & 1 \\ a_{21}-a_{22} & a_{22}-a_{23}\cdots\cdots a_{2,n-1}-a_{2n} & 1 \\ \cdots\cdots & \cdots\cdots & \\ a_{n1}-a_{n2} & a_{n2}-a_{n3}\cdots\cdots a_{n,n-1}-a_{nn} & 1 \end{vmatrix} = \sum_{i,j=1}^{n} A_{ij},$$

其中A_{ij}的含义同第4题。

6. 设$A=(a_{ij})_{n\times n}$是正交阵，求证：只要有某一个$a_{ij}\neq 0$使$a_{ij}+\left|\hat{A}_c\begin{pmatrix} i \\ j \end{pmatrix}\right|=0$，则对任何$1\leqslant i_1<i_2<\cdots<i_k\leqslant n$，$1\leqslant j_1<j_2<\cdots<j_k\leqslant n$，必有

$$\left|A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|+\left|\hat{A}_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right|=0.$$

7. 设n阶反对称阵A的顺序主子式$\Delta_{2k}\neq 0$，$\Delta_{2k-2}\neq 0$，这里$k\geqslant 2$，求证：$\Delta_{2k}\cdot\Delta_{2k-2}>0$。

§7 Cauchy-Binet 公式

本节应用行列式的第一降阶定理以及 Laplace 定理推导出重要的 Cauchy-Binet 公式。

定理 7.1 设 A 是 $n\times m$ 阵，B 是 $m\times n$ 阵，则

(i) 当 $n>m$ 时，恒有$|AB|=0$，

(ii) 当 $n\leqslant m$ 时，成立

$$|AB| = \sum_{1\leqslant j_1<\cdots<j_n\leqslant m} \left|A\begin{pmatrix} 1 & 2\cdots n \\ j_1 & j_2\cdots j_n \end{pmatrix}\right|\cdot\left|B\begin{pmatrix} j_1 & j_2\cdots j_n \\ 1 & 2\cdots n \end{pmatrix}\right|. \quad (1)$$

【证明】 由行列式的升阶公式可知

$$|AB| = |0-(-A)I_m^{-1}B| = \frac{1}{|I_m|}\begin{vmatrix} I_m & B \\ -A & 0 \end{vmatrix} = \begin{vmatrix} I_m & B \\ -A & 0 \end{vmatrix}. \quad (2)$$

当 $n>m$ 时，把(2)式右边的行列式按它的最后 n 列展开，则在它的 $m+n$ 行中任取 n 行所组成的每个 n 阶子式中至少有 $n-m$ 个行是零向量，所以由 Laplace 定理可知，这个行列式为零，这就证明了(i)。

今证(ii)。当 $m\geqslant n$ 时，仍把(2)式右边的 $m+n$ 阶行列式按它的最后 n 列展开，则 $\begin{pmatrix}B\\0\end{pmatrix}$ 中不等于零的 n 阶子式最多有 C_m^n 个，而其中的任意一个都是某个

$$\left|B\begin{pmatrix}j_1 & j_2\cdots j_n\\1 & 2\cdots n\end{pmatrix}\right|,\quad 1\leqslant j_1<j_2<\cdots<j_n\leqslant m.$$

故若能证明每个 $\left|B\begin{pmatrix}j_1 & j_2\cdots j_n\\1 & 2\cdots n\end{pmatrix}\right|$ 在这个 $m+n$ 阶行列式中的代数余子式恰好是对应的 $\left|A\begin{pmatrix}1 & 2\cdots n\\j_1 & j_2\cdots j_n\end{pmatrix}\right|$，则由 Laplace 定理，立即得到(1)式。

现在就来证明上述这个事实。先求出特殊的子式 $\left|B\begin{pmatrix}1 & 2\cdots n\\1 & 2\cdots n\end{pmatrix}\right|$ 的代数余子式。把上述 $m+n$ 阶行列式作如下分块

$$\begin{vmatrix}I_m & B\\-A & 0\end{vmatrix}=\begin{matrix}n\\m-n\\n\end{matrix}\begin{vmatrix}\overset{n}{I_n} & \overset{m-n}{0} & \overset{n}{B_1}\\0 & I_{m-n} & B_2\\-A_1 & -A_2 & 0\end{vmatrix},\tag{3}$$

其中

$$A_1=A\begin{pmatrix}1 & 2\cdots n\\1 & 2\cdots n\end{pmatrix},\quad B_1=B\begin{pmatrix}1 & 2\cdots n\\1 & 2\cdots n\end{pmatrix}.$$

易见 $|B_1|$ 的代数余子式是

$$(-1)^{\sum_{k=1}^{n}(k+(m+k))}\begin{vmatrix}0 & I_{m-n}\\-A_1 & -A_2\end{vmatrix}$$

$$=(-1)^{mn}\begin{vmatrix}0 & I_{m-n}\\-A_1 & -A_2\end{vmatrix}.\tag{4}$$

把(4)式的最右边的行列式按它的前 n 列展开，得到

$$\begin{vmatrix} 0 & I_{m-n} \\ -A_1 & -A_2 \end{vmatrix} = (-1)^{\sum_{k=1}^{n}((m-n+k)+k)} \cdot |-A_1| \cdot |I_{m-n}|$$

$$= (-1)^{n(m-n+1)}|A_1|. \tag{5}$$

把(5)式代入(4)式，可知 $|B_1|$ 的代数余子式恰好是

$$(-1)^{mn}(-1)^{n(m-n+1)}|A_1| = |A_1| = \left| A\begin{pmatrix} 1 & 2\cdots n \\ 1 & 2\cdots n \end{pmatrix} \right|.$$

今再求任一 n 阶子式 $\left| B\begin{pmatrix} j_1 & j_2\cdots j_n \\ 1 & 2\cdots n \end{pmatrix} \right|$ 的代数余子式。把行列式

$$\Delta = \begin{vmatrix} I_m & B \\ -A & 0 \end{vmatrix}$$

的子块 $(I_m \quad B)$ 中 B 的第 $j_1, j_2, \cdots, j_n$ 行依次调到第 $1, 2, \cdots, n$ 行的位置，同时，把 Δ 的子块 $\begin{pmatrix} I_m \\ -A \end{pmatrix}$ 中 A 的第 $j_1, j_2, \cdots, j_n$ 列依次调到第 $1, 2, \cdots, n$ 列的位置。易见，这样得到的新行列式 $\mathbf{D}$ 恰好等于 D，且

$$\mathbf{D} = \begin{vmatrix} I_n & 0 & B\begin{pmatrix} j_1 & j_2\cdots j_n \\ 1 & 2\cdots n \end{pmatrix} \\ 0 & I_{m-n} & B_3 \\ -A\begin{pmatrix} 1 & 2\cdots n \\ j_1 & j_2\cdots j_n \end{pmatrix} & -A_3 & 0 \end{vmatrix},$$

这又化到刚才讨论过的情形，故可知 $\left| B\begin{pmatrix} j_1 & j_2\cdots j_n \\ 1 & 2\cdots n \end{pmatrix} \right|$ 在 $\mathbf{D}(=\Delta)$ 中的代数余子式就是 $A\begin{pmatrix} 1 & 2\cdots n \\ j_1 & j_2\cdots j_n \end{pmatrix}$。 证毕。

例 1 由定理 7.1 的(i)可知，当 $n>2$ 时，n 阶行列式

$$\begin{vmatrix} a_1c_1+b_1d_1 & a_1c_2+b_1d_2\cdots\cdots a_1c_n+b_1d_n \\ a_2c_1+b_2d_1 & a_2c_2+b_2d_2\cdots\cdots a_2c_n+b_2d_n \\ \cdots\cdots & \cdots\cdots \\ a_nc_1+b_nd_1 & a_nc_2+b_nd_2\cdots\cdots a_nc_n+b_nd_n \end{vmatrix}$$

$$
=\left|\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix}\begin{pmatrix} c_1 & c_2\cdots c_n \\ d_1 & d_2\cdots d_n \end{pmatrix}\right| = 0.
$$

例 2（Lagrange 恒等式） 设 a_i, b_i 都是复数，$i=1,2,\cdots,m$, $m\geqslant 2$，求证

$$
\left(\sum_{i=1}^{m}|a_i|^2\right)\left(\sum_{i=1}^{m}|b_i|^2\right)-\left|\sum_{i=1}^{m}a_i\bar{b}_i\right|^2
= \sum_{1\leqslant i<j\leqslant m}|a_ib_j-a_jb_i|^2. \tag{6}
$$

【证明】

$$
\left(\sum_{i=1}^{m}|a_i|^2\right)\left(\sum_{i=1}^{m}|b_i|^2\right)-\left|\sum_{i=1}^{m}a_i\bar{b}_i\right|^2
$$

$$
=\begin{vmatrix} \sum\limits_{i=1}^{m}|a_i|^2 & \sum\limits_{i=1}^{m}a_i\bar{b}_i \\ \overline{\left(\sum\limits_{i=1}^{m}a_i\bar{b}_i\right)} & \sum\limits_{i=1}^{m}|b_i|^2 \end{vmatrix}
$$

$$
=\left|\begin{pmatrix} a_1a_2\cdots a_m \\ b_1b_2\cdots b_m \end{pmatrix}\begin{pmatrix} \bar{a}_1 & \bar{b}_1 \\ \bar{a}_2 & \bar{b}_2 \\ \vdots & \vdots \\ \bar{a}_m & \bar{b}_m \end{pmatrix}\right|.
$$

由假设 $m\geqslant 2$，故由 Cauchy-Binet 公式，(7)式化为

$$
\left(\sum_{i=1}^{m}|a_i|^2\right)\left(\sum_{i=1}^{m}|b_i|^2\right)-\left|\sum_{i=1}^{m}a_i\bar{b}_i\right|^2
$$

$$
=\sum_{1\leqslant i<j\leqslant m}\begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix}\begin{vmatrix} \bar{a}_i & \bar{b}_i \\ \bar{a}_j & \bar{b}_j \end{vmatrix}
$$

$$
=\sum_{1\leqslant i<j\leqslant m}|a_ib_j-a_jb_i|^2.
$$

证毕。

由于 Lagrange 恒等式的右端恒为非负，故立刻得到下面的 Cauchy-Schwarz 不等式

$$
\left|\sum_{i=1}^{m}a_i\bar{b}_i\right|^2\leqslant\left(\sum_{i=1}^{m}|a_i|^2\right)\left(\sum_{i=1}^{m}|b_i|^2\right). \tag{8}
$$

定理 7.2 设 $C=AB$，而 A 是 $m\times n$ 阵，B 是 $n\times l$ 阵，则

$$\left|C\begin{pmatrix} i_1 i_2\cdots i_k \\ j_1 j_2\cdots j_k\end{pmatrix}\right|$$

$$=\begin{cases}\sum\limits_{1\leqslant s_1<\cdots<s_k\leqslant n}\left|A\begin{pmatrix} i_1 i_2\cdots i_k \\ s_1 s_2\cdots s_k\end{pmatrix}\right|\cdot\left|B\begin{pmatrix} s_1 & s_2\cdots s_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right|, & \text{当 } k\leqslant n, \\ 0, & \text{当 } k>n。\end{cases} \tag{9}$$

【证明】 把A按它的行分块，B按它的列分块：

$$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix},\quad B=(\beta_1,\ \beta_2,\ \cdots,\ \beta_l),$$

则

$$\left|C\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right|=\left|\begin{pmatrix}\alpha_{i_1}\\ \alpha_{i_2}\\ \vdots\\ \alpha_{i_k}\end{pmatrix}(\beta_{j_1},\ \beta_{j_2},\ \cdots,\ \beta_{j_k})\right|$$

的右边的行列式中，前一个矩阵是 $k\times n$ 阵，后一个是 $n\times k$ 阵，故由定理 7.1，当 $k>n$ 时，$\left|C\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k\end{pmatrix}\right|=0$；当 $k\leqslant n$ 时，即得(9)式。 证毕。

例 3 设A是n阶非异实方阵，则 $A'A$ 的 n 个顺序主子式全大于零。

【证明】 任意取定k，$1\leqslant k\leqslant n$。由于 $|A|\neq 0$，A 的前 k 列中至少有某个 k 阶子式 $\left|A\begin{pmatrix} s_1 & s_2\cdots s_k \\ 1 & 2\cdots k\end{pmatrix}\right|\neq 0$，故由定理 7.2 可得

$$\left|(A'A)\begin{pmatrix} 1 & 2\cdots k \\ 1 & 2\cdots k\end{pmatrix}\right|$$

$$=\sum_{1\leqslant s_1<\cdots<s_k\leqslant n}\left|A'\begin{pmatrix} 1 & 2\cdots k \\ s_1 & s_2\cdots s_k\end{pmatrix}\right|\cdot\left|A\begin{pmatrix} s_1 & s_2\cdots s_k \\ 1 & 2\cdots k\end{pmatrix}\right|$$

$$=\sum_{1\leqslant s_1<\cdots<s_k\leqslant n}\left|A\begin{pmatrix} s_1 & s_2\cdots s_k \\ 1 & 2\cdots k\end{pmatrix}\right|^2>0。$$

习　　题

1. 设A是$n\times m$复矩阵，求证$|\bar{A}'A|\geqslant 0$，而不等号成立的充分必要条件是：$n\geqslant m$，且A中至少有一个m阶子式不为零。

2. 设a_i, b_i, c_i, d_i都是复数，$i=1, 2, \cdots, n$。求证

$$\left(\sum_{i=1}^{n} a_i\bar{c}_i\right)\left(\sum_{i=1}^{n} b_i\bar{d}_i\right)-\left(\sum_{i=1}^{n} a_i\bar{d}_i\right)\left(\sum_{i=1}^{n} b_i\bar{c}_i\right)$$
$$=\sum_{1\leqslant i<j\leqslant n}(a_ib_j-a_jb_i)\overline{(c_id_j-c_jd_i)}.$$

3. 设

$$V=\begin{pmatrix}1 & 1\cdots\cdots 1\\ a_1 & a_2\cdots\cdots a_m\\ \cdots\cdots & \cdots\cdots\cdots\\ a_1^n & a_2^n\cdots\cdots a_m^n\end{pmatrix}_{(n+1)\times m},$$

其中$a_1, a_2, \cdots, a_m$是两两互异的非零实数，且$n\leqslant m$。求证：VV'的任何主子式全大于零。

4. 设S是n阶非零反对称阵，求证

$$\begin{vmatrix}2I_n & S\\ S & 2I_n\end{vmatrix}>2^{2n}.$$

选　做　题

1. 设$a_i, b_i, i=1, 2, 3, 4$是八个实数。

(i) 试由

$$\begin{vmatrix}a_1 & a_2\\ -a_2 & a_1\end{vmatrix}=a_1^2+a_2^2,$$

证明$(a_1^2+a_2^2)(b_1^2+b_2^2)=c_1^2+c_2^2$，这里$c_1, c_2$是由$a_1, a_2, b_1, b_2$经过加、减、乘三种运算得出的数。

(ii) 记$x_1=a_1+\sqrt{-1}\,a_3$，$x_2=a_2+\sqrt{-1}\,a_4$，试由等式

$$\begin{vmatrix}x_1 & x_2\\ -\bar{x}_2 & \bar{x}_1\end{vmatrix}=|x_1|^2+|x_2|^2,$$

证明

$$\left(\sum_{i=1}^{4} a_i^2\right)\left(\sum_{i=1}^{4} b_i^2\right)=\sum_{i=1}^{4} c_i^2,$$

这里c_i是由a_j, b_k经过加、减、乘三种运算而得到的数。

（注：这是“平方和问题”的一部分，该问题的提法是：任意两个具有2^k个

实数的平方和的乘积是否仍为 2^k 个实数的平方和? $k=0$ 的情形显然是对的,本题是说当 $k=1, 2$ 时,回答都是肯定的,对 $k=3$, 要利用所谓 Cayley-Dickson 八元数,也可给出肯定的回答,但对 $k\geqslant 4$ 的情形, Hurwitz 在 1933 年已把它否定了。

2. (i) 设

$$A=\begin{pmatrix} a+b\sqrt{-1} & -1 \\ 1 & a-b\sqrt{-1} \end{pmatrix},$$

这里 a 与 b 为实数. 求证: A 的任一元素与它的代数余子式的共轭复数相等.

(ii) 设 $A=(a_{ij})$ 是 n 阶非零复阵, A_{ij} 是 a_{ij} 的代数余子式, 如果 $a_{ij}=\overline{A_{ij}}$, $i, j=1, 2, \cdots, n$, 求证: A 是非异阵,且 $\overline{|A|}=|A|^{n-1}$.

3. 设 n 是任意正整数。

(i) 记 $f(n)=1+2+\cdots+n$。如果设 $f(n)=an^2+bn+c$, 则用 $f(n+1)-f(n)=n+1$ 证明 a 与 b 满足如下的上三角形方程组

$$\begin{cases} b+a=1 \\ 2a=1, \end{cases}$$

试由此求出 $f(n)$ 的值。

(ii) 设 $f(n)=\sum_{l=1}^{n} l^2=an^3+bn^2+cn+d$。 仿照 (i) 中方法求出 a, b, c 所满足的上三角形方程组, 并由此证明

$$f(n)=\frac{1}{6}n(n+1)(2n+1)。$$

(iii) 用上述方法求

$$f(n)=1\cdot 2+2\cdot 3+\cdots+(n-1)\cdot n$$

的值。

(iv) 设

$$f(n)=\sum_{l=1}^{n} l^k=a_{k+1}n^{k+1}+a_k n^k+\cdots+a_1 n+a_0,$$

写出 a_1, a_2, $\cdots$, a_{k+1} 所满足的上三角形方程组。

4. 设 λ_1, λ_2, $\cdots$, λ_n 是 $x^n-1=0$ 的 n 个根,求证

$$\begin{vmatrix} \lambda_1 & a & \cdots\cdots & a \\ a & \lambda_2 & \cdots\cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \cdots\cdots & \lambda_n \end{vmatrix}=(-1)^{n-1}\{(n-1)a^n+1\}。$$

5. 称矩阵 A 为**整数阵**,如果 A 中任一元素都是整数。行列式为 ± 1 的方阵称为**么模阵**。

(i) 求证么模整数阵的逆阵仍是么模整数阵;

(ii) 设 A 是 n 阶整数方阵，问对任何 n 维整数列向量 b，n 阶线性方程组 $Ax=b$ 必有整数解的充分必要条件是什么？并证明之。

6. 把下面的行列式化成上三角形行列式，从而求值：

$$
\begin{vmatrix}
a_1 & a_2 & a_3 & \cdots & a_{n-1} & a_n \\
-y_1 & x_1 & & & & \\
 & -y_2 & x_2 & & & \\
 & & \ddots & \ddots & & \\
 & & & & x_{n-2} & \\
 & & & & -y_{n-1} & x_{n-1}
\end{vmatrix}.
$$

(注：把某些行列式化为三角形行列式，是计算行列式的第一个方法，若能把某些行列式先化为形如

$$
\begin{vmatrix}
* & * & \cdots & * \\
* & * & & \\
* & & \ddots & \\
\vdots & & & \\
* & & & *
\end{vmatrix}
\quad 或 \quad
\begin{vmatrix}
* & * & * & \cdots & * & * \\
* & * & & & & \\
 & * & * & & & \\
 & & \ddots & \ddots & * & \\
 & & & & * & *
\end{vmatrix}
$$

的行列式，再分别仿照 §2 例 7 和本题的方法求值，往往是有效的。)

7. 证明

$$
\begin{vmatrix}
1 & 2 & 3 \cdots\cdots n-1 & n \\
2 & 3 & 4 \cdots\cdots n & 1 \\
3 & 4 & 5 \cdots\cdots 1 & 2 \\
\cdots & \cdots & \cdots\cdots\cdots & \cdots \\
n & 1 & 2 \cdots\cdots n-2 & n-1
\end{vmatrix}
=(-1)^{\frac{n(n-1)}{2}}\cdot\frac{n(n+1)}{2}\cdot n^{n-2}.
$$

8. 用找递推关系式的方法证明如下的 **Cauchy 行列式**

$$
D_n=\begin{vmatrix}
\frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} & \cdots\cdots & \frac{1}{x_1+y_n} \\
\frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} & \cdots\cdots & \frac{1}{x_2+y_n} \\
\cdots & \cdots & \cdots & \cdots \\
\frac{1}{x_n+y_1} & \frac{1}{x_n+y_2} & \cdots\cdots & \frac{1}{x_n+y_n}
\end{vmatrix}
=\frac{\prod\limits_{1\le i<j\le n}(x_i-x_j)(y_i-y_j)}{\prod\limits_{i,j=1}^{n}(x_i+y_j)}.
$$

（注：找 D_{n-1} 与 D_{n-2} 或 D_n 与 D_{n-1}、D_{n-2} 的递推关系，从而算出 D_n，是计算行列式的第二个方法。）

9. 计算下面的 n 阶行列式

$$\begin{vmatrix} x & y & y & \cdots\cdots & y \\ z & x & y & \cdots\cdots & y \\ z & z & x & \cdots\cdots & y \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z & z & z & \cdots\cdots & x \end{vmatrix}.$$

10. 计算下面的循环行列式

$$D=\begin{vmatrix} c_1 & c_2 & c_3\cdots\cdots c_{n-1} & c_n \\ c_n & c_1 & c_2\cdots\cdots c_{n-2} & c_{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ c_2 & c_3 & c_4\cdots\cdots c_n & c_1 \end{vmatrix}.$$

（提示：以 Vandermonde 行列式 $V=\prod\limits_{1\leqslant j<i\leqslant n}(\xi_i-\xi_j)$ 乘 D，算得

$$VD=V\cdot\prod_{i=1}^{n}(c_1+\xi_ic_2+\xi_i^2c_3+\cdots+\xi_i^{n-1}c_n).$$

从而算出 D。这里，$\xi_1,\xi_2,\cdots,\xi_n$ 是 n 个不同的 n 次单位根，即 $\xi_i^n=1$.）

（注：用行列式乘法规则是计算行列式的第三个方法。）

11. 计算 n 阶行列式

$$\begin{vmatrix} x_1 & a & a\cdots\cdots a & a \\ a & x_2 & a\cdots\cdots a & a \\ \cdots & \cdots & \cdots & \cdots \\ a & a & a\cdots\cdots x_{n-1} & a \\ a & a & a\cdots\cdots a & x_n \end{vmatrix}.$$

（提示：把原行列式升阶得

$$D=\begin{vmatrix} 1 & a & a\cdots\cdots a & a \\ 0 & x_1 & a\cdots\cdots a & a \\ 0 & a & x_2\cdots\cdots a & a \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a & a\cdots\cdots x_{n-1} & a \\ 0 & a & a\cdots\cdots a & x_n \end{vmatrix}$$

$$
= \begin{vmatrix} 1 & a & a & \cdots\cdots & a & a \\ -1 & x_1-a & 0 & \cdots\cdots & 0 & 0 \\ -1 & 0 & x_2-a & \cdots\cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & 0 & 0 & \cdots\cdots & x_{n-1}-a & 0 \\ -1 & 0 & 0 & \cdots\cdots & 0 & x_n-a \end{vmatrix},
$$

而右端的 $n+1$ 阶行列式是容易计算的。)

12．计算下面的 n 阶行列式

$$
D = \begin{vmatrix} 1 & x_1 \cdots\cdots x_1^{k-1} & x_1^{k+1} \cdots\cdots x_1^n \\ 1 & x_2 \cdots\cdots x_2^{k-1} & x_2^{k+1} \cdots\cdots x_2^n \\ \cdots & \cdots & \cdots \\ 1 & x_n \cdots\cdots x_n^{k-1} & x_n^{k+1} \cdots\cdots x_n^n \end{vmatrix}.
$$

(提示：把 D 升阶，先算出 $n+1$ 阶 Vandermonde 行列式

$$
\Delta = \begin{vmatrix} 1 & x \cdots\cdots x^{k-1} & x^k & x^{k+1} \cdots\cdots x^n \\ 1 & x_1 \cdots\cdots x_1^{k-1} & x_1^k & x_1^{k+1} \cdots\cdots x_1^n \\ 1 & x_2 \cdots\cdots x_2^{k-1} & x_2^k & x_2^{k+1} \cdots\cdots x_2^n \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_n \cdots\cdots x_n^{k-1} & x_n^k & x_n^{k+1} \cdots\cdots x_n^n \end{vmatrix}
$$

而 Δ 中元素 x^k 的余子式就是 D。)

(注：把原行列式 Δ 升阶，算出这个行列式后，从而算出 Δ，这是计算行列式的第四种方法。)

13．设 A，B 都是 n 阶阵。

(i) 从 A 中划掉第 i 行，余下的 $(n-1)\times n$ 阵记为 $\tilde{A}$，从 B 中划掉第 j 列，余下的 $n\times(n-1)$ 阵记为 $\tilde{B}$，$1\leqslant i,\ j\leqslant n$。求证

$$
|\tilde{A}\tilde{B}| = (-1)^{i+j}\sum_{k=1}^{n}\left|\hat{A}_c\begin{pmatrix} i \\ k \end{pmatrix}\right|\cdot\left|\hat{B}_c\begin{pmatrix} k \\ j \end{pmatrix}\right|;
$$

(ii) 求证

$$
\mathrm{adj}(AB) = (\mathrm{adj}\,B)(\mathrm{adj}\,A).
$$

14．(Jacobi) 设 A 是 n 阶非异阵，则对任何 k，$1\leqslant k\leqslant n-1$，必成立

$$
\left|(\mathrm{adj}\,A)\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right| = |A|^{k-1}\left|\hat{A}_c\begin{pmatrix} j_1 & j_2\cdots j_k \\ i_1 & i_2\cdots i_k \end{pmatrix}\right|.
$$

(提示：用行列式升阶公式，并用 $\mathrm{adj}\,A = A^{-1}|A|$.)

(注：在第三章中将证，上述等式对奇异阵 A 也正确.)

15．设 n 阶阵 A 的 k 阶顺序主子式 $\Delta_k=\left|A\begin{pmatrix}1 & 2\cdots k\\ 1 & 2\cdots k\end{pmatrix}\right|\neq 0,\quad 1\leqslant k\leqslant n-1$，记

$$D_{ij}=\left|A\begin{pmatrix}1 & 2\cdots k & i\\ 1 & 2\cdots k & j\end{pmatrix}\right|,\ i,\ j=k+1,\ k+2,\ \cdots,\ n.$$

求证

$$|A|=\frac{1}{\Delta_k^{n-k-1}}\begin{vmatrix}D_{k+1,k+1} & D_{k+1,k+2}\cdots\cdots D_{k+1,n}\\ D_{k+2,k+1} & D_{k+2,k+2}\cdots\cdots D_{k+2,n}\\ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots\\ D_{n,k+1} & D_{n,k+2}\ \cdots\cdots\ D_{nn}\end{vmatrix}.$$

称它为行列式的**第三降阶定理**。

（提示：把 A 作如下分块

$$A=\begin{pmatrix}A\begin{pmatrix}1 & 2\cdots k\\ 1 & 2\cdots k\end{pmatrix} & A_1\\ A_2 & A\begin{pmatrix}k+1 & k+2\cdots n\\ k+1 & k+2\cdots n\end{pmatrix}\end{pmatrix},$$

再对 $|A|$ 应用第一降阶定理. 记

$$A\begin{pmatrix}k+1 & k+2\cdots n\\ k+1 & k+2\cdots n\end{pmatrix}-A\left(A\begin{pmatrix}1 & 2\cdots k\\ 1 & 2\cdots k\end{pmatrix}\right)^{-1}A_1$$

$$=\begin{pmatrix}c_{k+1,k+1}\ \cdots\cdots\ c_{k+1,n}\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ c_{n,k+1},\ \cdots\cdots\ c_{nn}\end{pmatrix}.$$

再对 c_{ij} 用行列式的升阶公式。$i,\ j=k+1,\ \cdots,\ n$.）

16．记

$$S_{ij}=\left|A\begin{pmatrix}1\ 2\cdots k\ i\\ 1\ 2\cdots k\ j\end{pmatrix}\right|,\ k+1\leqslant i,\ j\leqslant n,$$

称 $n-k$ 阶阵

$$S=\begin{pmatrix}s_{k+1,k+1}\ \cdots\ s_{k+1,n}\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ s_{n,k+1}\ \cdots\cdots\ s_{nn}\end{pmatrix}=S\begin{pmatrix}k+1,\ \cdots n\\ k+1,\ \cdots n\end{pmatrix}$$

为 **Sylvester 矩阵**。如果 n 阶阵 A 的 k 阶顺序主子式 $\Delta_k\neq 0$，求证

$$|S|=\Delta_k^{n-k-1}|A|.$$

我们把此一等式称为 **Sylvester 恒等式**。

（提示：用第三降阶定理）

第三章　线性代数方程组与矩阵的秩

本章完整地解决了一般的线性方程组的求解问题。讨论了线性方程组在行列式与矩阵论上的应用，并对解决线性方程组求解问题的核心概念——矩阵的秩作了较为细致的讨论。

§1　向量组的线性无关与矩阵的秩

一、方程组求解与向量组的线性关系

在理论问题中以及应用上，常需考虑含 m 个方程、n 个未知量（数）的一般线性代数方程组：

$$\sum_{j=1}^{n} a_{ij}x_j = b_i,\ i=1,\ 2,\ \cdots,\ m, \tag{1}$$

为简便计，称(1)为 $m\times n$ **方程组**，它也可写成“**矩阵方程**”的形状

$$Ax=b, \tag{2}$$

其中 $A=(a_{ij})_{m\times n}$ 称为(1)的**系数矩阵**，而称 $\tilde{A}=(A,\ b)$ 为 A 的**增广矩阵**，列向量 $x=(x_1,\ x_2,\ \cdots,\ x_n)'$ 称为**未知向量**，而 $b=(b_1,\ b_2,\ \cdots,\ b_m)'$ 称为**常（数）向量**。

某个列向量 $(k_1,\ k_2,\ \cdots,\ k_n)'$ 称为(1)的解（**向量**），如果 $x_j=k_j$ $(j=1,\ 2,\ \cdots,\ n)$，代入(1)后使之成为恒等式。设 $l\times n$ 方程组

$$Bx=d \tag{3}$$

的任一解是(2)的解，且(2)的任一解也是(3)的解，则称(2)与(3)**同解**，或称(3)是(2)的**同解方程组**。我们总设法把(2)化成一个形状简单的同解方程组，从而求出解 x。

$m\times n$ 方程组要解决的问题是，怎样判断它是否有解？在有解时如何求出它的解？解的形状怎样？解有什么性质？它有哪些应用？

为解决上述有关线性(代数)方程组的求解问题，先由下例观察一下，应引进怎样的“工具”？

例1 在 5×3 方程组：

$$\begin{cases} x_1+x_2+x_3=3 \\ 2x_1+3x_2+4x_3=9 \\ 3x_1+3x_2-5x_3=1 \\ x_1+2x_2+3x_3=6 \\ x_1+3x_2+13x_3=17 \end{cases} \tag{4}$$

中，易见第二个方程减去第一个方程就是第四个方程，第二个方程的两倍减去第三个方程就是第五个方程，这说明如下的 3×3 方程组：

$$\begin{cases} x_1+x_2+x_3=3 \\ 2x_1+3x_2+4x_3=9 \\ 3x_1+3x_2-5x_3=1 \end{cases} \tag{5}$$

是原 5×3 方程组 (4) 的同解方程组，或者说，(4)中的最后两个方程对前三个方程来说是多余的，可把它们舍掉。

我们注意到以下事实：在由(4)确定(5)的过程中，实际上并没有涉及未知数 x_1、x_2、x_3，而是把一个方程作为一个整体，进行加、减与数乘运算的，因此，上述过程完全可用对增广矩阵的各个行向量之间作相应的运算来描述，确切地说，设

$$\tilde{A}=(A,\ b)=\begin{pmatrix} 1 & 1 & 1 & 3 \\ 2 & 3 & 4 & 9 \\ 3 & 3 & -5 & 1 \\ 1 & 2 & 3 & 6 \\ 1 & 3 & 13 & 17 \end{pmatrix}=\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{pmatrix}$$

则以下两个等式：

$$\alpha_4=\alpha_2-\alpha_1,\ \alpha_5=2\alpha_2-\alpha_3 \tag{6}$$

分别反映了第四个方程对于第一、二两个方程，第五个方程对于第二、三两个方程是多余的这一事实，因此(4)中有否多余方程的问题就归结为增广矩阵的行向量之间是否有形如(6)的那种关系式。

出于这种考虑，我们引进：

定义 设 $\alpha, \alpha_1, \alpha_2, \cdots, \alpha_s$ 是 n 维行向量，如果存在数 $k_1, k_2, \cdots, k_s$，使

$$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \tag{7}$$

则称 α 是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的**线性组合**，或称 α 可由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ **线性表出**，那些 k_i 称为**表出系数**。

由于(6)式可写成：

$$\alpha_4 = (-1)\alpha_1 + 1\cdot\alpha_2 + 0\cdot\alpha_3,\ \alpha_5 = 0\cdot\alpha_1 + 2\alpha_2 + (-1)\alpha_3 \tag{8}$$

故 α_4 与 α_5 都是 $\alpha_1, \alpha_2, \alpha_3$ 的线性组合，即(4)的最后两个方程对前三个方程来说是多余的。

若把(8)式改写成：

$$\begin{aligned} 1\cdot\alpha_1 + (-1)\alpha_2 + 0\cdot\alpha_3 + 1\cdot\alpha_4 + 0\cdot\alpha_5 = 0, \\ 0\cdot\alpha_1 + (-2)\alpha_2 + 1\cdot\alpha_3 + 0\cdot\alpha_4 + 1\cdot\alpha_5 = 0, \end{aligned} \tag{9}$$

则上两式也反映了(4)中有多余方程这一事实，于是一般地引进：

定义 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 都是 n 维行向量，如果存在不全为零的数 $k_1, k_2, \cdots, k_s$，使

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0 \tag{10}$$

则称 $\alpha_1, \alpha_2, \cdots, \alpha_s$ **线性相关**。否则，称 $\alpha_1, \alpha_2, \cdots, \alpha_s$ **线性无关**，或**线性独立**，也就是说，(10)式成立当且仅当 $k_i = 0,\ i = 1, 2, \cdots, s$。

注意：关于线性相关定义中的条件“不全为零”是必需的，否则的话，任意向量组都满足(10)式了。而在关于线性表出的定义中，那些表出系数 k_i 却可以全为零，事实上，零向量是任意向量组的线性组合。

由(8)式与(9)式可看出，向量组的线性相关与线性组合有着密切的联系。事实上有下面的

命题 1 $s(\geqslant 2)$ 个 n 维行向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关的充要条件是，至少有某个 α_i 是其余向量的线性组合。

【证明】 如果 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关，则在(10)式中至少有某个系数，例如 $k_i \neq 0$，于是(10)式可改写为

$$\alpha_i=\left(-\frac{k_1}{k_i}\right)\alpha_1+\cdots+\left(-\frac{k_{i-1}}{k_{i+1}}\right)\alpha_{i-1}+\left(\frac{-k_{i+1}}{k_i}\right)\alpha_{i+1}$$
$$+\cdots+\left(-\frac{k_s}{k_i}\right)\alpha_s$$

这说明 α_i 是 $\alpha_1, \cdots, \alpha_{i-1}, \alpha_{i+1}, \cdots, \alpha_s$ 的线性组合。

反之，设

$$\alpha_i=l_1\alpha_1+\cdots+l_{i-1}\alpha_{i-1}+l_{i+1}\alpha_{i+1}+\cdots+l_s\alpha_s$$

则

$$l_1\alpha_1+\cdots+l_{i-1}\alpha_{i-1}+(-1)\alpha_i+l_{i+1}\alpha_{i+1}+\cdots+l_s\alpha_s$$

由于 $l_1, \cdots, l_{i-1}, -1, l_{i+1}, \cdots, l_s$ 不全为零，故 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关。　　　　证毕。

由命题 1 可知，如果 $\tilde{A}=(A, b)$的 n 个行向量线性相关，则在 $Ax=b$ 中必有多余方程，反之亦然。

继续考察例 1，很自然地要问，在 3×3 方程组(5)中是否还有多余方程？如果还有，则以下三个 4 维行向量：

$\alpha_1=(1, 1, 1, 3)$, $\alpha_2=(2, 3, 4, 9)$, $\alpha_3=(3, 3, -5, 1)$

应线性相关，也即下面的 4×3 方程组：

$$\begin{cases} k_1+2k_2+3k_3=0 \\ k_1+3k_2+3k_3=0 \\ k_1+4k_2-5k_3=0 \\ 3k_1+9k_2+k_3=0 \end{cases}$$

应有非零解，这是不可能的，因为由 Cramer 法则，上方程组的前三个方程，由于系数行列式不为零，只有零解，也就是 $\alpha_1, \alpha_2, \alpha_3$ 线性无关，所以在(4)的前三个方程中再也没有多余的方程了，此时，称这三个方程所成的方程组(5)为(4)的极大无关（独立）方程组，相应地，称向量组 $\alpha_1, \alpha_2, \alpha_3$ 为 $\tilde{A}$ 的行向量组 $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ 的极大线性无关(独立)组，它有这样两个性质：$\alpha_1, \alpha_2, \alpha_3$ 线性无关；$\tilde{A}$ 的向量都是 $\alpha_1, \alpha_2, \alpha_3$ 的线性组合(由 (8) 式可知 α_4, α_5 是 $\alpha_1, \alpha_2, \alpha_3$ 的线性组合，又 $\alpha_1, \alpha_2, \alpha_3$ 的每一个显然是 $\alpha_1, \alpha_2, \alpha_3$ 的线性组合)，于是一般地引进：

定义　如果在 n 维行向量组 $\alpha_1, \alpha_2, \cdots, \alpha_5$ 中有“部分向量

组"$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$使

(i) $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 线性无关；

(ii) 每一 α_j 都是 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 的线性组合，$1 \leqslant j \leqslant s$，则称 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 为 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的**极大线性无关（独立）组**，而称 r 为行向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的**秩（数）**。

相应地，任一 $m \times n$ 阵 A 的 m 个行向量中的极大线性无关向量的个数（即行向量的秩）称为 A 的**行秩**。

由此可知，如果 $\widetilde{A}$ 的极大线性无关行向量组是 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$，则 $r \times n$ 方程组：

$$\begin{cases} a_{i_1 1}x_1 + a_{i_1 2}x_2 + \cdots + a_{i_1 n}x_n = b_{i_1} \\ a_{i_2 1}x_1 + a_{i_2 2}x_2 + \cdots + a_{i_2 n}x_n = b_{i_2} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i_r 1}x_1 + a_{i_r 2}x_2 + \cdots + a_{i_r n}x_n = b_{i_r} \end{cases} \tag{11}$$

就是 $A\underset{\sim}{x} = b$ 的极大无关（独立）方程组，它与 $Ax = b$ 同解，因此，如能找出 $\widetilde{A}$ 的行向量组的极大线性无关组，也就得到了同解方程组 (11)，故只要把 (11) 的有关解的问题搞清楚，也就解决了原方程组 $Ax = b$ 的求解问题。

如果对列向量组类似地引进线性相关、线性无关、线性组合的概念，则命题 1 相应地对列向量组也成立，且可得

命题 2 如把 $m \times n$ 阵 A 按它的列分块：$A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$，则 $m \times n$ 方程组 $Ax = b$ 有解的充要条件为，b 是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合。

【证明】 如果 $x = (k_1, k_2, \cdots, k_n)'$ 是 $Ax = b$ 的解，则

$$(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = b,$$

故 $b = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$。

反之，把上述证明过程倒推回去，即知 $(k_1, k_2, \cdots, k_n)$ 是 $Ax = b$ 的解。

证毕。

综上所述，为要找 $Ax=b$ 的极大独立方程组，就必须讨论行向量组的线性无关(相关)性以及矩阵的行秩；为要讨论方程组是否有解，就必须考虑列向量组的线性无关(相关)性。所以，方程组的求解问题与向量组的线性关系式是密切相联的，以下将用后者作为“工具”去解决前者的问题。

二、向量组无关性判别法则及其应用

本段如无特别说明，所说向量既可以是行向量，也可以是列向量。

首先，由线性相关的定义可知，单独一个向量 α 线性无关的充要条件是，α 为非零向量。其次，向量组有下列两个常用性质：

命题 3 设 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 与 $\alpha_1, \alpha_2, \cdots, \alpha_p, \alpha_{p+1}, \cdots, \alpha_q$ 是 n 维向量组，如果前者(线性)相关，则后者也相关(部分相关，整体也相关)。换言之，若后者(线性)无关，则前者也无关。

【证明】 设 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性相关，则存在不全为零的数 $k_1, k_2, \cdots, k_p$，使

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_p\alpha_p=0,$$

于是由

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_p\alpha_p+0\cdot\alpha_{p+1}+\cdots+0\cdot\alpha_q=0,$$

就知 $\alpha_1, \alpha_2, \cdots, \alpha_p, \alpha_{p+1}, \cdots, \alpha_q$ 线性无关（因为 $k_1, k_2, \cdots, k_p, 0, \cdots, 0$ 仍然不全为零）。 证毕。

设 $\alpha=(a_1, a_2, \cdots, a_r)$为 r 维行向量，称 $r+s$ 维行向量：$(a_1, a_2, \cdots, a_r, b_1, \cdots, b_s)$为 α 的**接长向量**。同样可定义列向量的接长向量。

命题 4 线性无关向量组的接长向量组也线性无关。

【证明】 就行向量组来证明，列向量组仿此，设 $\beta_i=(\alpha_i, \delta_i)$ 是 α_i 的接长向量，$i=1, 2, \cdots, t$，且 $\alpha_1, \alpha_2, \cdots, \alpha_t$ 线性无关，则由

$$k_1(\alpha_1, \delta_1)+k_2(\alpha_2, \delta_2)+\cdots+k_t(\alpha_t, \delta_t)=(0, 0),$$

可得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_t\alpha_t=0,$$

所以 $k_1=k_2=\cdots=k_t=0$，即 $\beta_1, \beta_2, \cdots, \beta_t$ 线性无关。　　**证毕。**

下面的定理是本章以后一些定理的基础。

定理 1.1　(基本定理 1)设

$$\alpha_i=(a_{i1}, a_{i2}, \cdots, a_{in}), i=1, 2, \cdots, p,$$

$p\leqslant n$，则 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性无关的充要条件是，在 $p\times n$ 阵：

$$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12}\cdots a_{1n}\\ a_{21} & a_{22}\cdots a_{2n}\\ \cdots & \cdots\\ a_{p1} & a_{p2}\cdots a_{pn}\end{pmatrix}$$

中至少有一个 p 阶子式不等于零。

【证明】　必要性。对 p 用归纳法。

当 $p=1$ 时，$A=\alpha_1$ 为线性无关向量，所以 $\alpha_1\neq0$，它自然有某个分量(即 A 的一阶子式)不等于零。

设当向量个数小于 p 时，结论正确。今由假设，$\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性无关，故由命题 3 知，$\alpha_1, \alpha_2, \cdots, \alpha_{p-1}$ 也线性无关，于是由归纳法假设，如下 $(p-1)\times n$ 阵：

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_{p-1}\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{p-1,1} & a_{p-1,2} & \cdots & a_{p-1,n}\end{pmatrix}$$

中必有 $p-1$ 阶非零子式，为书写简便起见，不妨设 $p-1$ 阶顺序主子式：

$$D=\left|A\begin{pmatrix}1 & 2\cdots p-1\\ 1 & 2\cdots p-1\end{pmatrix}\right|\neq0.$$

今考虑 A 的如下的 $n-(p-1)$ 个 p 阶子式：

$$D_j=\left|A\begin{pmatrix}1 & 2\cdots p-1 & p\\ 1 & 2\cdots p-1 & j\end{pmatrix}\right|=\begin{vmatrix}a_{11} & \cdots & a_{1,p-1} & a_{1j}\\ \cdots & \cdots & \cdots & \cdots\\ a_{p-1,1} & \cdots & a_{p-1,p-1} & a_{p-1,j}\\ a_{p1} & \cdots & a_{p,p-1} & a_{pj}\end{vmatrix}$$

$j=p, p+1, \cdots, n$，则 D_j 中的(i, p)元 a_{ij} 在 D_j 中的代数余子式都(可记)为 A_{ip}，$i=1, 2, \cdots, p$。容易看出，对所有 $j=p, p+1, \cdots, n$ 来说，A_{ip} 仅与 i 有关而与 j 无关，且 $D=A_{pp}$。

对于 p 阶行列式 D_p，由行列式的性质可知，

$$\begin{cases} a_{11}A_{1p}+a_{21}A_{2p}+\cdots+a_{p1}A_{pp}=0 \\ a_{12}A_{1p}+a_{22}A_{2p}+\cdots+a_{p2}A_{pp}=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1,p-1}A_{1p}+a_{2,p-1}A_{2p}+\cdots+a_{p,p-1}A_{pp}=0 \end{cases} \tag{12}$$

再将 D_p、D_{p+1}、$\cdots$、D_n 都按它的第 p 列展开，可得

$$\begin{cases} a_{1p}A_{1p}+a_{2p}A_{2p}+\cdots+a_{pp}A_{pp}=D_p \\ a_{1,p+1}A_{1p}+a_{2,p+1}A_{2p}+\cdots+a_{p,p+1}A_{pp}=D_{p+1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1n}A_{1p}+a_{2n}A_{2p}+\cdots+a_{pn}A_{pp}=D_n \end{cases} \tag{13}$$

如果 $D_p=D_{p+1}=\cdots=D_n=0$，则(13)诸等式右端全为零，于是连同(12)中的 $p-1$ 个等于零的等式，可合并写成行向量等式：

$$A_{1p}\alpha_1+A_{2p}\alpha_2+\cdots+A_{p-1,p}\alpha_{p-1}+D\alpha_p=0$$

因 $D\neq 0$，故 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性相关，这与假设矛盾，所以 D_p，$D_{p+1}, \cdots, D_n$ 这些 p 阶子式中至少有一个不等于零。

充分性。不失一般性，可设 A 的 p 阶顺序主子式不等于零，于是

$$A_p=A\begin{pmatrix}1 & 2\cdots p \\ 1 & 2\cdots p\end{pmatrix}$$

是非异阵，把 A_p 按它的行分块：

$$A_p=\begin{pmatrix}\beta_1 \\ \beta_2 \\ \vdots \\ \beta_p\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12}\cdots a_{1p} \\ a_{21} & a_{22}\cdots a_{2p} \\ \cdots\cdots & \cdots\cdots \\ a_{p1} & a_{p2}\cdots a_{pp}\end{pmatrix},$$

今证 $\beta_1, \beta_2, \cdots, \beta_p$ 线性无关。事实上，由

$$0=k_1\beta_1+k_2\beta_2+\cdots+k_p\beta_p=(k_1, k_2, \cdots, k_p)A_p$$

可得

$$(k_1, k_2, \cdots, k_p) = 0 \cdot A_p^{-1} = 0,$$

故 $\beta_1, \beta_2, \cdots, \beta_p$ 线性无关，因 α_i 是 βi 的接长向量，$i=1, 2, \cdots, p$，故由命题 4 知 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 也线性无关。 证毕。

例 2 由于矩阵

$$\begin{pmatrix} 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -3 \end{pmatrix}$$

的 2 阶子式 $\begin{vmatrix} 1 & 2 \\ 0 & -3 \end{vmatrix} \neq 0$，故向量 $(0, 1, -1, 2)$，$(0, 0, 0, -3)$ 必线性无关。

例 3 对任意一个 $p \times n$ 阵：

$$A = \begin{pmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{21} & a_{22} \cdots a_{2n} \\ \cdots\cdots & \cdots\cdots \\ a_{p1} & a_{p2} \cdots a_{pn} \end{pmatrix},$$

$p \leqslant n$，必存在主对角线上只出现 1 或 -1 的 $p \times n$ 对角形阵 D，使 $A+D$ 的 p 个行向量线性无关。

【证明】 显然，$a_{11}+1$ 与 $a_{11}-1$ 这两个数中至少有一个不为零，不妨设 $a_{11}-1 \neq 0$，则

$$\alpha_1 = (a_{11}-1, a_{12}, \cdots, a_{1n}) \neq 0,$$

故 α_1 线性无关，再考虑向量 α_1 与向量

$$\beta_1 = (a_{21}, a_{22}+1, a_{23}, \cdots, a_{2n})$$

$$\beta_2 = (a_{21}, a_{22}-1, a_{23}, \cdots, a_{2n}),$$

则可证

$$\Delta_1 = \begin{vmatrix} a_{11}-1 & a_{12} \\ a_{21} & a_{22}+1 \end{vmatrix}, \qquad \Delta_2 = \begin{vmatrix} a_{11}-1 & a_{12} \\ a_{21} & a_{22}-1 \end{vmatrix}$$

中至少有一个不为零，不然的话，经过简单计算即可得出 $a_{11}-1=0$ 的矛盾，故不妨设 $\Delta_1 \neq 0$。由定理 1.1 可知，α_1, β_1 必线性无关，于是找到了 $2 \times n$ 对角形阵：

$$D_2 = \begin{pmatrix} -1 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 \cdots 0 \end{pmatrix}$$

使
$$\begin{pmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \end{pmatrix} + \begin{pmatrix} -1 & 0\cdots 0 \\ 0 & 1\cdots 0 \end{pmatrix}$$
的两个行向量线性无关。

归纳地，设对 A 的 $p-1$ 阶顺序主子阵 $A_{p-1}=(a_{ij})_{(p-1)\times(p-1)}$ 已找到了 $(p-1)\times n$ 阵。

$$\begin{pmatrix} d_1 & & & \\ & d_2 & & 0 \\ & & \ddots & \\ & & & d_{p-1} \end{pmatrix} = (D_{p-1}, 0),$$

其中 $D_{p-1}=[d_1, d_2, \cdots, d_{p-1}]$，$\alpha_i = 1$ 或 -1，$1 \leqslant i \leqslant p$，使得

$$|A_{p-1}+D_{p-1}| = \begin{vmatrix} a_{11}+d_1 & a_{12} & \cdots a_{1,p-1} \\ a_{21} & a_{22}+d_2 & \cdots a_{2,p-1} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{p-1,1} & a_{p-1,2} & \cdots a_{p-1,p-1}+d_{p-1} \end{vmatrix} \neq 0,$$

则同样可证

$$\Delta_3 = \begin{vmatrix} A_{p-1}+D_{p-1} & u \\ v & a_{pp}+1 \end{vmatrix}$$
$$\Delta_4 = \begin{vmatrix} A_{p-1}+D_{p-1} & u \\ v & a_{pp}-1 \end{vmatrix}$$

中至少有一个不等于零，此处，

$$u=(a_{1p}, a_{2p}, \cdots, a_{p-1,p})', \ v=(a_{p1}, a_{p2}, \cdots, a_{p,p-1}),$$

事实上，如果 $\Delta_3=\Delta_4=0$，则因 $A_{p-1}+D_{p-1}$ 是非异阵，故由行列式的第一降阶定理可知，

$$\Delta_3 = |A_{p-1}+D_{p-1}|(a_{pp}+1-v(A_{p-1}+D_{p-1})^{-1}u)=0,$$
$$\Delta_4 = |A_{p-1}+D_{p-1}|(a_{pp}-1-v(A_{p-1}+D_{p-1})^{-1}u)=0,$$

再由 $|A_{p-1}+D_{p-1}| \neq 0$ 及上两式立即得出 $1=-1$ 的矛盾，故又找到了 $p\times n$ 阵：

$$D = \begin{pmatrix} D_{p-1} & 0 & \\ 0 & d_p & 0 \end{pmatrix}, \ d_p = 1 \text{ 或 } d_p = -1$$

使 $A+D$ 的 p 阶顺序主子式(即 Δ_3 或 Δ_4)不为零，所以由定理 1.1 知 $A+D$ 的 p 个行向量线性无关。　　证毕。

如果类似地引进列向量组的极大线性无关组（这只要把行向量组的极大线性无关组的定义中的"行"字换成"列"字），并相应地称任一 $m\times n$ 阵的列向量组的极大线性无关组中的向量个数为矩阵的**列秩**，则可得

推论 1 n 阶非异阵 A 的行秩与列秩都是 n. 反之，若 n 阶阵 A 的行秩（或列秩）为 n，则 A 是非异阵。

这是明显的，因为由定理 1.1 可知，上述结论对行向量组是正确的，再由 $|A'|=|A|$ 就可推知，此推论对 A' 的行向量组即 A 的列向量组来说也正确。

由于推论 1，故也称非异阵为**满秩阵**。

推论 2 任何 $n+1$ 个 n 维向量必线性相关。因而向量个数大于向量维数的向量组必是线性相关组。

【证明】 设 $\alpha_1, \alpha_2, \cdots, \alpha_n, \alpha_{n+1}$ 是 n 维行向量组，如果 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性相关，则由命题 3，$\alpha_1, \alpha_2, \cdots, \alpha_n, \alpha_{n+1}$ 也线性相关；如果 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关，则由推论 1，方阵 $(\alpha_1', \alpha_2', \cdots, \alpha_n')$ 为非异阵，故由 Cramér 法则，方程组：

$$(\alpha_1', \alpha_2' \cdots, \alpha_n')\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\alpha_{n+1}'$$

有唯一解：$(x_1, x_2, \cdots, x_n)'=(k_1, k_2, \cdots, k_n)'$，于是得到

$$\alpha_{n+1}'=k_1\alpha_1'+k_2\alpha_2'+\cdots+k_n\alpha_n',$$

或者

$$\alpha_{n+1}=k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n,$$

再由命题 1，$\alpha_1, \alpha_2, \cdots, \alpha_n, \alpha_{n+1}$ 线性相关。

对列向量组，可类似地证明。 证毕。

定理 1.2（基本定理 2） 设 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 与 $\beta_1, \beta_2, \cdots, \beta_q$ 是两个 n 维向量组，如果

(i) 每个 β_i 都是 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 的线性组合，$1\leqslant i\leqslant q$；

(ii) $q>p$，

则 $\beta_1, \beta_2, \cdots, \beta_q$ 必线性相关。

【证明】 就行向量的情形证明,列向量可仿此证出(或在行向量的情形已证好的基础上，应用它们的转置向量有相同性质这一点也立即可得)。

由假设(i)依次可得

$$\beta_1=(c_{11}, c_{12}, \cdots, c_{1p})\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{pmatrix},$$

$$\beta_2=(c_{21}, c_{22}, \cdots, c_{2p})\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{pmatrix},$$

$$\cdots\cdots$$

$$\beta_q=(c_{q1}, c_{q2}, \cdots, c_{qp})\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{pmatrix},$$

上面诸式可合并成矩阵等式:

$$\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_q\end{pmatrix}=\begin{pmatrix}c_{11} & c_{12}\cdots c_{1p}\\ c_{21} & c_{22}\cdots c_{2p}\\ \cdots & \cdots\cdots\\ c_{q1} & c_{q2}\cdots c_{qp}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{pmatrix}=C\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_p\end{pmatrix}, \tag{14}$$

其中 $C=(c_{ij})_{q\times p}$。再由假设 (ii), $q>p$, 故由推论 2 可知, C 的 q 个行向量必线性相关,也即存在$(k_1, k_2, \cdots, k_q)\neq 0$, 使

$$(k_1, k_2, \cdots k_q)C=0。$$

于是在(14)式两端左乘$(k_1, k_2, \cdots, k_q)$就得

$$k_1\beta_1+k_2\beta_2+\cdots+k_q\beta_q=0,$$

所以 $\beta_1, \beta_2, \cdots, \beta_q$ 线性相关。 证毕。

由定理 1.2 显然可得下面三个推论:

推论 1 如果向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的极大线性无关组中向量

个数是 r, $r<s$, 则 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 中任意$r+1$ 个向量必线性相关。

推论 2 设 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 与 $\beta_1, \beta_2, \cdots, \beta_q$ 是两个 n 维向量组，如果：

(i) 每个 β_i 都是 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 的线性组合，$1\leqslant i\leqslant q$；

(ii) $\beta_1, \beta_2, \cdots, \beta_q$ 线性无关，

则 $q\leqslant p$。

论推 3 设 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 与 $\beta_1, \beta_2, \cdots, \beta_q$ 是两个线性无关向量组，如果每个 β_i 都是 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 的线性组合，$1\leqslant i\leqslant q$；每个 α_j 都是 $\beta_1, \beta_2, \cdots, \beta_q$ 的线性组合，则 $p=q$。

一个向量组中的极大线性无关组可能不止一组，例如第一段例 1 的 A 的行向量 $\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5$ 中，$\alpha_1, \alpha_2, \alpha_3$ 与 $\alpha_1, \alpha_2, \alpha_5$ 都是极大无关组。然而，这两个极大无关组中向量的个数却是相同的，事实上有下述一般结论。

定理 1.3 一个向量组中任意两个极大无关组所含的向量个数相同。

【证明】 设 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 与 $\beta_1, \beta_2, \cdots, \beta_t$ 是同一个向量组中的两个极大线性无关组，则每一 α_i 可由 $\beta_1, \beta_2, \cdots, \beta_t$ 线性表出，$1\leqslant i\leqslant r$；每一 β_j 可由 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性表出，$1\leqslant j\leqslant t$，故由定理 1.2 的推论 3，$r=t$。 证毕。

由定理 1.3 显然可知下面推论的正确性：

推论 任一 $m\times n$ 阵 A 的行(列)秩由 A 唯一确定。

三、矩阵的秩的基本性质

定理 1.4 $m\times n$ 阵 A 中有 p 个行向量线性无关当且仅当 A 中有 p 个列向量线性无关。

【证明】 设

$$A=\begin{pmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots & \cdots \\ a_{m1} & a_{m2}\cdots a_{mn} \end{pmatrix}=\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}=(\beta_1, \beta_2, \cdots, \beta_n),$$

这里 α_i 与 β_j 分别是 A 的第 i 行与第 j 列，$i=1,2,\cdots,m$；$j=1,2,\cdots,n$。

设 $\alpha_{i_1},\alpha_{i_2},\cdots,\alpha_{i_p}$ 线性无关，则由定理 1.1 可知，在由这 p 个行所成的 $p\times n$ 阵中必有 p 阶非零子式，设为

$$\left|A\begin{pmatrix}i_1 & i_2\cdots i_p\\ j_1 & j_2\cdots j_p\end{pmatrix}\right|=\begin{vmatrix}a_{i_1j_1} & a_{i_1j_2}\cdots a_{i_1j_p}\\ a_{i_2j_1} & a_{i_2j_2}\cdots a_{i_2j_p}\\ \cdots\cdots & \cdots\cdots\cdots\\ a_{i_pj_1} & a_{i_pj_2}\cdots a_{i_pj_p}\end{vmatrix}\neq 0,$$

于是由定理 1.1 的推论 1，

$$A\begin{pmatrix}i_1 & i_2\cdots i_p\\ j_1 & j_2\cdots j_p\end{pmatrix}$$

的 p 个列向量线性无关，而 $\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_p}$ 分别是这 p 个 p 维向量的接长向量，故由命题 4，它们也线性无关。

反之，设 A 有 p 个列向量线性无关，则 A' 有 p 个行向量线性无关，据刚才必要性的证明知道，A' 应有 p 个列向量线性无关，也即 A 有 p 个行向量线性无关。　　证毕。

命题 5　设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha$ 线性相关，但 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关，则 α 必是 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的线性组合。

【证明】　由假设，存在不全为零的数 $k_1,k_2,\cdots,k_r,k$，使

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r+k\alpha=0, \tag{15}$$

如果 $k=0$，则由(15)式以及 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关的假设可知，$k_1=k_2=\cdots=k_r=0$，此为矛盾，故 $k\neq0$。于是(15)式可改写为

$$\alpha=\left(-\frac{k_1}{k}\right)\alpha_1+\left(\frac{-k_2}{k}\right)\alpha_2+\cdots+\left(\frac{-k_r}{k}\right)\alpha_r,$$

这说明 α 是 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的线性组合。　　证毕。

定理 1.5　任一 $m\times n$ 阵 A 的行秩等于它的列秩。

【证明】　设 A 的行秩为 r，则 A 有 r 个行向量构成极大线性无关组，且由定理 1.2 的推论 1 可知，A 的任意 $r+1$ 个行向量(如果有的话，即 $r<m$) 必线性相关，于是由定理 1.4，A 有 r 个列向量 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_r}$ 线性无关，则 A 的任意 $r+1$ 个列向量(如果

有的话)必线性相关,故对任一列向量 α_j, $j \neq j_1, j_2, \cdots, j_r$, α_j, $\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$ 也线性相关,于是由命题 5, α_j 是 $\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$ 的线性组合,又每一 α_{j_k} 显然是 $\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$ 的线性组合, $1 \leqslant k \leqslant r$, 所以 $\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$ 是 A 的列向量组的极大线性无关组,于是 A 的列秩也是 r。 证毕。

由于定理 1.5, 今后我们就称 A 的行(列)秩为 A 的**秩**,并记为 $\mathrm{r}(A)$。又由于 A' 的行秩就是 A 的列秩,而 A 的行(列)秩由 A 唯一确定(定理 1.3 的推论),故定理 1.5 的另一说法是, A 与 A' 的秩相同,即

$$\mathrm{r}(A) = \mathrm{r}(A'). \tag{16}$$

命题 6 如果向量组 $\delta_1, \delta_2, \cdots, \delta_r$ 的每个向量 δ_i 都是向量 $\beta_1, \beta_2, \cdots, \beta_s$ 的线性组合,而每个 β_j 又是向量 $\alpha_1, \alpha_2, \cdots, \alpha_t$ 的线性组合,则每个 δ_i 必是 $\alpha_1, \alpha_2, \cdots, \alpha_t$ 的线性组合, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant s$。

【证明】 就列向量组来证明。由假设,

$$\delta_i = k_{i1}\beta_1 + k_{i2}\beta_2 + \cdots + k_{is}\beta_s, \quad 1 \leqslant i \leqslant r$$

以及

$$\beta_j = (\alpha_1, \alpha_2, \cdots, \alpha_t)\begin{pmatrix} l_{j1} \\ l_{j2} \\ \vdots \\ l_{jt} \end{pmatrix}, \quad j = 1, 2, \cdots, s,$$

故

$$\delta_i = k_{i1}(\alpha_1, \alpha_2, \cdots, \alpha_t)\begin{pmatrix} l_{11} \\ l_{12} \\ \vdots \\ l_{1t} \end{pmatrix}$$

$$+ k_{i2}(\alpha_1, \alpha_2, \cdots, \alpha_t)\begin{pmatrix} l_{21} \\ l_{22} \\ \vdots \\ l_{2t} \end{pmatrix}$$

$$+\cdots+k_{is}(\alpha_1,\alpha_2,\cdots,\alpha_t)\begin{pmatrix}l_{s1}\\l_{s2}\\\vdots\\l_{st}\end{pmatrix}$$

$$=(\alpha_1,\alpha_2,\cdots,\alpha_t)\begin{pmatrix}\sum_{j=1}^{s}k_{ij}l_{j1}\\\sum_{j=1}^{s}k_{ij}l_{j2}\\\vdots\\\sum_{j=1}^{s}k_{ij}l_{jt}\end{pmatrix},$$

即 δ_i 是 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 的线性组合,行向量组可仿此证得。证毕。

例 4 设 $A=(a_{ij})_{m\times n}$ 的秩为 r，且 A 的 i_1、i_2、$\cdots i_r$ 行与 A 的 j_1、j_2、$\cdots$、j_r 列分别是 A 的行向量组与列向量组的极大线性无关组,求证:

$$\left|A\begin{pmatrix}i_1 & i_2\cdots i_r\\j_1 & j_2\cdots j_r\end{pmatrix}\right|\neq 0.$$

【证明】 把 A 按它的行与列分块:

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1j_1} & \cdots & a_{1j_2} & \cdots & a_{1j_r} & \cdots & a_{1n}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ a_{i_11} & \cdots & a_{i_1j_1} & \cdots & a_{i_1j_2} & \cdots & a_{i_1j_r} & \cdots & a_{i_1n}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ a_{i_21} & \cdots & a_{i_2j_1} & \cdots & a_{i_2j_2} & \cdots & a_{i_2j_r} & \cdots & a_{i_2n}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ a_{i_r1} & \cdots & a_{i_rj_1} & \cdots & a_{i_rj_2} & \cdots & a_{i_rj_r} & \cdots & a_{i_rn}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ a_{m1} & \cdots & a_{mj_1} & \cdots & a_{mj_2} & \cdots & a_{mj_r} & \cdots & a_{mn}\end{pmatrix}=\begin{pmatrix}\alpha_1\\\vdots\\\alpha_{i_1}\\\vdots\\\alpha_{i_2}\\\vdots\\\alpha_{i_r}\\\vdots\\\alpha_m\end{pmatrix}$$

$$=(\beta_1,\cdots,\beta_{j_1},\cdots,\beta_{j_2},\cdots,\beta_{j_r},\cdots,\beta_n)$$

作 $m\times r$ 阵

$$A_1=\begin{pmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_r} \end{pmatrix}=(\beta_{j1},\ \beta_{j2}\cdots,\ \beta_{jr})=\begin{pmatrix} \tilde{\alpha}_1 \\ \vdots \\ \tilde{\alpha}_{i_1} \\ \vdots \\ \tilde{\alpha}_{i_2} \\ \vdots \\ \alpha_{i_r} \\ \vdots \\ \alpha_m \end{pmatrix}$$

由于$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$是极大无关组，故任一$\alpha_i$是$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$的线性组合,因此任一$\tilde{\alpha}_i$自然也是$\tilde{\alpha}_{i_1}, \tilde{\alpha}_{i_2}, \cdots, \tilde{\alpha}_{i_r}$的线性组合,又由假设$\beta_{j_1},\beta_{j_2},\cdots,\beta_{j_r}$线性无关,故$A_1$的秩为$r$。今证$\tilde{\alpha}_{i_1}, \tilde{\alpha}_{i_2}, \cdots, \tilde{\alpha}_{i_r}$线性无关。如果$\tilde{\alpha}_{i_1}, \tilde{\alpha}_{i_2}, \cdots, \tilde{\alpha}_{i_r}$线性相关，并设$\tilde{\alpha}_1, \tilde{\alpha}_2, \cdots, \hat{\alpha}_t$是它的极大线性无关组,$t<r$,则任一$\hat{\alpha}_{ij}$是$\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \tilde{\alpha}_t$的线性组合,于是由命题 6, 任一$\tilde{\alpha}_i$也是$\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_t$的线性组合,这说明$A_1$的秩为$t$,这与秩的唯一性相矛盾(定理 **1.3** 的推论),所以$\tilde{\alpha}_{i_1}, \tilde{\alpha}_{i_2}, \cdots, \tilde{\alpha}_{i_r}$线性无关,由定理 1.1 的推论 **1**,

$$\left|A\begin{pmatrix} i_1 & i_2\cdots i_r \\ j_1 & j_2\cdots j_r \end{pmatrix}\right|\neq 0.$$

证毕。

由矩阵的秩的定义及命题 6 显然可得:

引理 把$m\times n$阵A按它的列分块: $A=(\alpha_1, \alpha_2, \cdots, \alpha_n)$, 如果$\alpha_1, \alpha_2, \cdots \alpha_n$中有某$n-r$个列，使其中的任一列都是($\alpha_1, \alpha_2, \cdots, \alpha_n$的)另外$r$个列的线性组合,则$\mathrm{r}(A)\leqslant r$。

定理 1.6 设A是$m\times n$阵,B是$n\times l$阵,则

$$\mathrm{r}(AB)\leqslant\min(\mathrm{r}(A),\ \mathrm{r}(B)).\tag{17}$$

【证明】 设$\mathrm{r}(B)=r$,则B有r个列线性无关,不失一般性,令它们是$\beta_1, \beta_2, \cdots, \beta_r$,则

$$\beta_j=k_{1j}\beta_1+k_{2j}\beta_2+\cdots+k_{rj}\beta_r,\ r+1\leqslant j\leqslant l,$$

所以

$$A\beta_j = k_{1j}A\beta_1 + k_{2j}A\beta_2 + \cdots + k_{rj}A\beta_r,\ r+1 \leqslant j \leqslant l, \tag{18}$$

又

$$AB = (A\beta_1, A\beta_2, \cdots, A\beta_r, A\beta_{r+1}, \cdots, A\beta_l), \tag{19}$$

(18)与(19)满足引理的条件，故得

$$\mathrm{r}(AB) \leqslant r = \mathrm{r}(B). \tag{20}$$

另外，由于任意矩阵与其转置矩阵有相同的秩(由(16)式)，且由(20)式可得

$$\mathrm{r}(AB) = \mathrm{r}((AB)') = \mathrm{r}(B'A') \leqslant \mathrm{r}(A') = \mathrm{r}(A), \tag{21}$$

(20)与(21)式说明(17)式是正确的。 证毕。

推论 设A是$m \times n$阵，P与Q分别是m阶与n阶非异阵，则

$$\mathrm{r}(AQ) = \mathrm{r}(PA) = \mathrm{r}(PAQ) = \mathrm{r}(A). \tag{22}$$

【证明】 令$PA = C$，则$A = P^{-1}C$，由(17)式，

$$\mathrm{r}(C) = \mathrm{r}(PA) \leqslant \mathrm{r}(A)$$

$$\mathrm{r}(A) = \mathrm{r}(P^{-1}C) \leqslant \mathrm{r}(C)$$

所以

$$\mathrm{r}(A) = \mathrm{r}(C) = \mathrm{r}(PA)$$

同理可证：$\mathrm{r}(A) = \mathrm{r}(AQ)$，于是

$$\mathrm{r}(PAQ) = \mathrm{r}((PA)Q) = \mathrm{r}(PA) = \mathrm{r}(A).$$ 证毕。

上面的推论说明，初等变换不改变矩阵的秩，由于任一$m \times n$阵A必可找到m阶非异阵P与n阶非异阵Q，使

$$PAQ = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}, \tag{23}$$

(第一章§5)，且s由A唯一确定，故由推论可知，A的秩是s，反之，有下面的

定理 1.7 （基本定理 3） 设$m \times n$阵A的秩为r，则必可找到m阶非异阵P与n阶非异阵Q，使

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{24}$$

【证明】 因为对任何A，必存在m阶非异阵P与n阶非异阵

Q，使成立(23)式，令 $r(A)=r$，故由(22)式

$$r=r(PAQ)=r\begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}=s,$$

所以(23)式可写成(24)式。 证毕。

上面的两小段讨论说明，只要用初等变换把 A 化为对角形 $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$，则这个 r 就是 A 的秩，我们已在第一章做过多次，不过那时未提"秩"这一名词罢了。

定义 设 A 与 B 都是 $m\times n$ 阵，如果存在 m 阶非异阵 P 与 n 阶非异阵 Q，使 $B=PAQ$，则称 A 与 B **相抵**，或称 A **相抵于** B。

推论 A 与 B 相抵的充要条件是，$r(A)=r(B)$。

【证明】 由定理 1.6 的推论可知，必要性是显然的，今证充分性。设 $r(A)=r$，则 $r(B)=r$，故由定理 1.7，存在非异阵 P，R 与 Q，T，使

$$PAQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}=RBT,$$

于是 $B=(R^{-1}P)A(QT^{-1})$，即 A 与 B 相抵。 证毕。

定理 1.7 的很多应用将在以后陆续给出，最后讨论行列式与矩阵的秩的关系。

定理 1.8 若 $m\times n$ 阵 A 的秩为 r，则 A 至少有一个 r 阶子式不等于零，而 A 的任何高于 r 阶的子式(如果还有的话)必为零。

【证明】 因为 $r(A)=r$，故 A 的 r 个线性无关的行向量所成的 $r\times n$ 阵至少有一个 r 阶子式不等于零(由定理 1.1)。又如果 A 中有某个 $r+1$ 阶子式不等于零，则它所在的 A 的 $r+1$ 个行向量必线性无关，这与 $r(A)=r$ 的假设相矛盾，故 A 的所有 $r+1$ 阶子式全等于零，于是由行列式的展开式可知，A 的任何高于 r 阶的子式也都等于零。 证毕。

定义 设 D 是 $m\times n$ 阵 A 的某个 r 阶子式，则 A 中任一包含 D(即以 D 为子式)的 $r+1$ 阶子式称为 D 的**加边子式**。

下面的结论提供了用行列式求矩阵的秩的方法

定理 1.9 果 $m\times$ 如阵 A 中有某个 r 阶子式 D 不等于零而，D 的所有加边子式全等于零，则 A 的秩为 r。

【证明】 为书写简洁起见，不妨设 A 的 r 阶顺序主子式

$$D=\begin{vmatrix} a_{11} & a_{12}\cdots a_{1r} \\ a_{21} & a_{22}\cdots a_{2r} \\ \cdots\cdots & \cdots\cdots \\ a_{r1} & a_{r2}\cdots a_{rr} \end{vmatrix}\neq 0$$

于是 A 的前 r 行线性无关，考虑如下的 $(r+1)\times n$ 阵

$$\tilde{A}_p=\begin{pmatrix} a_{11} & a_{12}\cdots a_{1r} & a_{1,r+1}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2r} & a_{2,r+1}\cdots a_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{r1} & a_{r2}\cdots a_{rr} & a_{r,r+1}\cdots a_{rn} \\ a_{p1} & a_{p2}\cdots a_{pr} & a_{p,r+1}\cdots a_{pn} \end{pmatrix}=\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_r \\ \alpha_p \end{pmatrix}, r+1\leqslant p\leqslant m,$$

这里 $\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_p$ 是 A_p 的行向量，再把 $\tilde{A}_p$ 按它的列分块：

$$\tilde{A}_p=(\beta_1, \beta_2, \cdots, \beta_r, \beta_{r+1}, \cdots, \beta_n),$$

此处 β_j 均为 $r+1$ 维列向量，$j=1, 2, \cdots, n$。由于 $D\neq 0$，故由定理 1.1，$\beta_1, \beta_2, \cdots, \beta_r$ 线性无关，但由假设，D 的加边子式全为零，所以对任一 s，$r+1\leqslant s\leqslant n$，恒有

$$|\beta_1, \beta_2, \cdots, \beta_r, \beta_s|=0,$$

所以 $\beta_1, \beta_2, \cdots, \beta_r, \beta_s$ 线性相关，故由命题 5，β_s 是 $\beta_1, \beta_2, \cdots, \beta_r$ 的线性组合，$r+1\leqslant s\leqslant n$，所以 $r(\tilde{A}_p)=r$。于是 $\tilde{A}_p$ 的第 p 行 α_p 必是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的线性组合，$r+1\leqslant p\leqslant m$，故对 A 来说，$\alpha_1, \alpha_2, \cdots, \alpha_r$ 是 A 的行向量组的极大线性无关组，所以 A 的秩是 r。

例 5 因为矩阵

$$A=\begin{pmatrix} -2 & -1 & -1 & 3 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 2 & 1 & 2 & -1 & 0 \end{pmatrix}$$

的 2 阶子式

$$D=\begin{vmatrix} -1 & -1 \\ 0 & 1 \end{vmatrix}\neq 0,$$

而 D 的所有加边子式，

$$\begin{vmatrix}-2 & -1 & -1\\ 0 & 0 & 1\\ 2 & 1 & 2\end{vmatrix}, \begin{vmatrix}-1 & -1 & 3\\ 0 & 1 & 2\\ 1 & 2 & -1\end{vmatrix}, \begin{vmatrix}-1 & -1 & 1\\ 0 & 1 & 1\\ 1 & 2 & 0\end{vmatrix}$$

全等于零，所以 $r(A)=2$。

由定理 1.8 与定理 1.9 显然可得

推论 矩阵A的秩就是A中非零子式的最高阶数。

我们也可用推论所述的结论作为矩阵的秩的定义。不过，这样的定义可能不易了解矩阵的秩的由来以及它与解线性方程组的关系。

习 题

1. 判断下列各向量组是否线性无关

(i) $\alpha_1=(1, 0, -1)$, $\alpha_2=(3, 1, 2)$, $\alpha_3=(-1, 1, 0)$;

(ii) $\alpha_1=(1, 2, 3, 4)$, $\alpha_2=(2, 3, 4, 1)$, $\alpha_3=(3, 4, 1, 2)$;

(iii) $\alpha_i=(0, \cdots, 0, \overset{i}{1}, 0, \cdots, 0, a_{i,r+1}, \cdots, a_{in})$, $i=1, 2, \cdots, r$;

(iv) $\alpha_i=(1, 1, \cdots, \overset{i}{1}, 0, \cdots, 0)$, $i=1, 2, \cdots, n$.

2. 确定数a，使下列各向量组线性无关

( ) $\alpha_i=(0, \cdots, 0, \overset{i}{1}, a, 0, \cdots, 0)$, $i=1, 2, \cdots, n-1$,

$\alpha_n=(a, 0, \cdots, 0, 1)$;

(ii) $\alpha_i=(1, \cdots, 1, \overset{i}{1-a}, 1, \cdots, 1)$, $i=1, 2, \cdots, n$.

3. 确定数 $a_1, a_2, \cdots, a_r$，使下列向量组线性无关

$\alpha_i=(1, a_i, a_i^2, \cdots, a_i^r, \cdots, a_i^n)$, $i=1, 2, \cdots, r$.

4. 把向量β表成向量 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的线性组合，其中

$\beta=(0, 0, 0, 1)$, $\alpha_1=(1, 1, 0, 1)$, $\alpha_2=(2, 1, 3, 1)$,

$\alpha_3=(1, 1, 0, 0)$, $\alpha_4=(0, 1, -1, -1)$.

5. 设 $\alpha_1\neq 0$，求证：$\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性相关的充要条件是，至少有某个 $\alpha_i(2\leqslant i\leqslant r)$ 可由 $\alpha_1, \alpha_2, \cdots, \alpha_{i-1}$ 线性表出。

6. 设β可由 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性表出，但不能由 $\alpha_1, \alpha_2, \cdots, \alpha_{r-1}$ 线性表出，证明 α_r 可由 $\alpha_1, \alpha_2, \cdots, \alpha_{r-1}, \beta$ 线性表出。

7. 设n维标准单位向量 $e_1, e_2, \cdots, e_n$ 可由 n 维(列)向量 $\alpha_1, \alpha_2, \cdots,$

α_n 线性表出，求证：$\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关。

8. 求证：n 维向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关的充要条件是，任何 n 维向量都是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合。

9. 设 β 可由 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性表出，求证：表出系数是唯一的充要条件是，$\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关。

10. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\beta_1, \beta_2, \cdots, \beta_n$ 是两个 m 维行向量组，且存在 n 阶方阵 C，使

$$\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_n\end{pmatrix}=C\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}$$

(i) 如果 $\beta_1, \beta_2, \cdots, \beta_n$ 线性无关，求证：$\alpha_1, \alpha_2, \cdots, \alpha_n$ 也线性无关。并问，反之是否成立？若不成立，试举出反例。

(ii) 如果 $\beta_1, \beta_2, \cdots, \beta_n$ 线性无关，证明每一 α_i 可由 $\beta_1, \beta_2, \cdots, \beta_n$ 线性表出，$1\leqslant i\leqslant n$。

11. 设 A 是 n 阶阵，求证：必存在主对角元是 1 或 -1 的 n 阶对角阵 D，使 $|DA+I_n|\neq 0$，此处 I_n 是 n 阶单位阵。

12. 求下列矩阵的秩

(i) $\begin{pmatrix}-2 & 1 & -1 & -3\\ 1 & -1 & 0 & 1\\ 0 & 3 & 3 & 3\end{pmatrix}$；　(ii) $\begin{pmatrix}0 & 1 & 1 & -1\\ 2 & 1 & 3 & 1\\ 3 & 6 & 3 & 0\\ 0 & -1 & -1 & 1\end{pmatrix}$

13. 试就 $a_1, a_2, \cdots, a_r$ 的各种不同情形讨论下列 $r\times n$ 阵 A 的秩。其中，

$$A=\begin{pmatrix}1 & a_1\cdots a_1^{n-1}\\ 1 & a_2\cdots a_2^{n-1}\\ \cdots\cdots & \cdots\cdots\\ 1 & a_r\cdots a_r^{n-1}\end{pmatrix}.$$

14. 设 A 与 B 分别是 m 阶阵，求证：

$$r\begin{pmatrix}A & 0\\ 0 & B\end{pmatrix}=r(A)+r(B).$$

15. 设 A 与 B 都是方阵（未必同阶），求证：

$$r\begin{pmatrix}A & C\\ 0 & B\end{pmatrix}\geqslant r(A)+r(B)$$

当 A（或 B）是非异阵时，则（上式的）等号成立。

16. 设A是$m\times n$阵，

(i) 求证：对任一$l\times n$阵B，恒有

$$\max(\mathrm{r}(A),\mathrm{r}(B))\leqslant \mathrm{r}\begin{pmatrix}A\\B\end{pmatrix}\leqslant \mathrm{r}(A)+\mathrm{r}(B);$$

(ii) 在A中任取s行构成$s\times n$阵B，求证：

$$\mathrm{r}(B)\geqslant \mathrm{r}(A)-(m-s);$$

(iii) 在A中任意划掉s行与t列，得到$(m-s)\times(n-t)$阵C，试求$\mathrm{r}(A)$与$\mathrm{r}(C)$所满足的不等式，并证明之。

17. 设A与B都是$m\times n$阵，求证：

$$\mathrm{r}(A+B)\leqslant \mathrm{r}(A)+\mathrm{r}(B)。$$

(提示：因为 $A+B=(I_m, I_m)\begin{pmatrix}A\\B\end{pmatrix}$，故应用定理**1.6**及**16**题 (**i**) 即可得证。)

18. 设A与B分别是$m\times n$阵与$n\times l$阵，应用第二章定理7.2证明：$\mathrm{r}(AB)\leqslant\min(\mathrm{r}(A),\mathrm{r}(B))$。

19. 设秩为r的n阶(实)对称阵A的第i_1、i_2、…、i_r行线性无关，证明：

$$\left|A\begin{pmatrix}i_1 & i_2\cdots i_r\\ i_1 & i_2\cdots i_r\end{pmatrix}\right|\neq 0.$$

换言之，可得这样的著名结论："秩为r的(实)对称阵至少有一个r阶主子式不等于零"。

(提示：由A的对称性及本节例4的结论显然可证得)

20. 求证：秩为$2k$的n阶反对称(实)阵至少有一个$2k$阶非零子式。

21. 设$A=(\alpha_1, \alpha_2, \cdots, \alpha_n)$是$m\times n$阵，且

$$\prod_{i=1}^{n} k_i\neq 0,$$

求证：$\mathrm{r}(\alpha_1, \alpha_2, \cdots, \alpha_n)=\mathrm{r}(k_1\alpha_1, k_2\alpha_2, \cdots, k_n\alpha_n)$。

22. 设$A=(a_{ij})_{n\times n}$，$B=(b_{ij})_{n\times n}$，且$b_{ij}=(-1)^{i+j}a_{ij}$，$i, j=1, 2, \cdots, n$，求证：$\mathrm{r}(A)=\mathrm{r}(B)$。

23. 求证：n阶阵A是奇异阵的充要条件是，A至少有一行(列)是其他各行(列)的线性组合，或者说，$\mathrm{r}(A)<n$。

24. 设A是n阶奇异阵，$n\geqslant 2$，求证：

(i) $\mathrm{r}(\mathrm{adj}A)\leqslant 1$(当$\mathrm{r}(A)<n-1$时，显然$\mathrm{r}(\mathrm{adj}A)=0$)；

(ii) $\left|(\mathrm{adj}A)\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right| = 0$, $2\leqslant k\leqslant n$; $1\leqslant i_1<\cdots<i_k\leqslant n$; $1\leqslant j_1<\cdots<j_k\leqslant n$.

25. 设A是n阶阵，$n>2$，求证：

$$\mathrm{adj}(\mathrm{adj}A)=|A|^{n-2}A.$$

§2 方程组的解法及应用

一、方程组相容性判别法则、相容方程组的解法

定义 如果$m\times n$方程组$Ax=b$有解，则称它是**相容的**，或称$Ax=b$是**相容方程组**；如果$Ax=b$无解，则称它**不相容**，或称$Ax=b$是**矛盾方程组**。

定理 2.1 (**Kronecker**) $m\times n$方程组$Ax=b$相容的充要条件是，$\mathrm{r}(A)=\mathrm{r}(\widetilde{A})$。其中$\widetilde{A}=(A, b)$。

【证明】 设$\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$是$A$的列向量组$\alpha_1, \alpha_2, \cdots\alpha_n$的极大线性无关组，如果$Ax=b$有解，则由命题 2，$b$是$\alpha_1, \alpha_2, \cdots, \alpha_n$的线性组合，又每一$\alpha_j$是$\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$的线性组合，故由命题 6，$b$是$\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$的线性组合，这说明$\widetilde{A}$的秩也是$r$，故$\mathrm{r}(A)=\mathrm{r}(\widetilde{A})$。

反之，如$\mathrm{r}(A)=\mathrm{r}(\widetilde{A})$，则$b$是$\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$的线性组合，故$b$自然也是$\alpha_1, \alpha_2, \cdots\alpha_n$的线性组合，所以由命题 3，$Ax=b$有解。 证毕。

定理 2.2 设$Ax=b$是相容方程组，$\mathrm{r}(A)=r$，则

(i) 当$r=n$时，$Ax=b$只有一个解；

(ii) 当$r<n$时，$Ax=b$有无穷多个解。

【证明】 设$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$是$A=(\alpha_{ij})_{m\times n}$的行向量组的极大线性无关组，因$Ax=b$是相容方程组，故由定理 2.1，$\mathrm{r}(A)=\mathrm{r}(\widetilde{A})$，所以$\widetilde{A}$的行向量$\widetilde{\alpha}_{i_1}, \widetilde{\alpha}_{i_2}, \cdots, \widetilde{\alpha}_{i_r}$也是$\widetilde{A}$的极大线性无关组，此处$\widetilde{\alpha}_{i_k}=(\alpha_{i_k}, b_{i_k})$，$k=1, 2, \cdots, r$。于是$Ax=b$与下面的$r\times n$方程组同解：

$$\begin{cases} a_{i_1 1}x_1 + a_{i_1 2}x_2 + \cdots + a_{i_1 n}x_n = b_{i_1} \\ a_{i_2 1}x_1 + a_{i_2 2}x_2 + \cdots + a_{i_2 n}x_n = b_{i_2} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i_r 1}x_1 + a_{i_r 2}x_2 + \cdots + a_{i_r n}x_n = b_{i_r} \end{cases} \quad (1)$$

设 A 的 r 阶子式

$$\left| A\begin{pmatrix} i_1 & i_2 \cdots i_r \\ j_1 & j_2 \cdots j_r \end{pmatrix} \right| \neq 0,$$

又设 $j_{r+1}, j_{r+2}, \cdots, j_n$ 是 $1,2,\cdots,n$ 这 n 个数中去掉 $j_1, j_2, \cdots, j_r$ 后余下的 $n-r$ 个数，则(1)可化为

$$\begin{cases} a_{i_1 j_1}x_{j_1} + a_{i_1 j_2}x_{j_2} + \cdots + a_{i_1 j_r}x_{j_r} = b_{i_1} - a_{i_1 j_{r+1}}x_{j_{r+1}} - \cdots - a_{i_1 j_n}x_{j_n} \\ a_{i_2 j_1}x_{j_1} + a_{i_2 j_2}x_{j_2} + \cdots + a_{i_2 j_r}x_{j_r} = b_{i_2} - a_{i_2 j_{r+1}}x_{j_{r+1}} - \cdots - a_{i_2 j_n}x_{j_n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i_r j_1}x_{j_1} + a_{i_r j_2}x_{j_2} + \cdots + a_{i_r j_r}x_{j_r} = b_{i_r} - a_{i_r j_{r+1}}x_{j_{r+1}} - \cdots - a_{i_r j_n}x_{j_n} \end{cases} \quad (2)$$

由方程组(2)并应用 Cramer 法则可知，当 $r=n$ 时，(2)有唯一解，也即 $Ax=b$ 有唯一解；当 $r<n$ 时，由于 $x_{r+1}, \cdots, x_n$ 可以取任何数，故方程组(2)，也即 $Ax=b$ 有无穷多个解。　　证毕。

定理 2.1 解决了方程组是否有解的问题；定理 2.2 则提供了方程组有解时的解法以及解的性质，并且由此可得出解的具体形状，故 §1 提出的关于方程组求解的五个问题的前四个至此已全部解决。

由定理 2.2 的(ii)显然可得下列两个常用结论：

推论 （Cramer 法则的逆定理）如果 n 阶方程组 $Ax=b$ 有唯一解，则 A 必是非异阵。

定理 2.3　A 是奇异阵的充要条件为 $Ax=0$ 有非零解。

二、方程组的解的应用

以下四个例子显示出方程组的解的理论在行列式与矩阵论上的应用。

例 1　设 A 与 B 分别是 $m\times n$ 阵与 $m\times l$ 阵，且 $\mathrm{r}(A)=\mathrm{r}(A,$

B)，则必存在 $n\times l$ 阵 C，使 $AC=B$(即"矩阵方程"$AX=B$ 有解)。

【证明】 把 B 按它的列分块：$B=(\beta_1, \beta_2, \cdots, \beta_l)$，由于 $r(A)=r(A, B)$，故 $r(A)=r(A, \beta_j)$，$j=1, 2, \cdots, l$，于是 $m\times n$ 方程组，$Ax=\beta_j$ 必有解：$x=\gamma_j$，$j=1, 2, \cdots, l$，即

$$A\gamma_j=\beta_j,\ j=1, 2, \cdots, l,$$

记 $C=(\gamma_1, \gamma_2, \cdots, \gamma_l)$，于是由上式即得：$AC=B$。 证毕。

例 2 设 α, β 都是 n 维非零列向量，A 是 n 阶非异阵，则 $(\beta' A^{-1}\alpha)A-\alpha\beta'$ 是奇异阵。

【证明】 因为 $\alpha\neq 0$，而 A 是非异阵，故 $A^{-1}\alpha\neq 0$，又因

$$((\beta' A^{-1}\alpha)A-\alpha\beta')(A^{-1}\alpha)=(\beta' A^{-1}\alpha)\alpha-\alpha(\beta' A^{-1}\alpha)=0,$$

这说明方程组$((\beta' A^{-1}\alpha)A-\alpha\beta')x=0$ 有非零解，故由定理 2.3，$(\beta' A^{-1}\alpha)A-\alpha\beta'$ 是奇异阵。 证毕。

注意：当 $\beta' A^{-1}\alpha\neq 0$ 时，由行列式的第二降阶定理也易证 $(\beta' A^{-1}\alpha)A-\alpha\beta'$ 的奇异性(当 $\beta' A^{-1}\alpha=0$，则 $-\alpha\beta'$ 显然是奇异阵)。

例 3 (**Lévy-Hadamard**)① 如果方阵 $A=(a_{ij})_{n\times n}$ 满足：

$$|a_{ii}|>\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|,\ i=1, 2, \cdots, n, \tag{3}$$

则 A 必是非异阵，称满足条件 (3) 的方阵为**严格对角占优阵**或 **Hadamard 矩阵**。

【证明】 若 A 是奇异阵，则由定理 2.3 知，$Ax=0$ 有非零解：$x=(\xi_1, \xi_2, \cdots, \xi_n)'$，此处 $\xi_1, \xi_2, \cdots, \xi_n$ 不全为零。于是

$$\sum_{j=1}^{n} a_{ij}\xi_j=0,\ i=1, 2, \cdots, n,$$

把上式改写成：

$$a_{ii}\xi_i=-\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}\xi_j,\ i=1, 2, \cdots, n,$$

上式两边取绝对值，可得

① 也有人称例 3 的结论为 Lévy-Desplanqué 定理。

$$|a_{ii}|\cdot|\xi_i| = \left|-\sum_{\substack{j=1\\j\neq i}}^{n} a_{ij}\xi_j\right| \leqslant \sum_{\substack{j=1\\j\neq i}}^{n} |a_{ij}|\,|\xi_j|,\ i=1,\ 2,\ \cdots,\ n. \quad (4)$$

记

$$|\xi_t| = \max\{|\xi_1|,\ |\xi_2|,\ \cdots,\ |\xi_n|\},$$

并在(4)中特别取第 t 个不等式,则得

$$|a_{tt}|\,|\xi_t| \leqslant \sum_{\substack{j=1\\j\neq t}}^{n} |a_{tj}|\,|\xi_j| \leqslant \left(\sum_{\substack{j=1\\j\neq t}}^{n} |a_{tj}|\right)|\xi_t|. \quad (5)$$

因为 $\xi_1, \xi_2, \cdots, \xi_n$ 不全为零,故必 $|\xi_t| \neq 0$, 于是(5)式可化为:

$$|a_{tt}| \leqslant \sum_{\substack{j=1\\j\neq t}}^{n} |a_{tj}|,$$

这与假设(3)相矛盾,所以 A 必是非异阵。 证毕。

例 4 设 A 是 n 阶反对称(实)阵,D 是 n 阶实对角阵, 且其主对角元全大于零,求证: $|A+D|>0$。

【证明】 先证 $|A+D|\neq 0$。如果 $|A+D|=0$, 则由定理 2.3, 存在 n 维实的列向量 $\xi\neq 0$, 使$(A+D)\xi=0$, 由于 A 是反对称阵, 故

$$\xi'(A+D)\xi = \xi'D\xi = 0.$$

记 $\xi=(\xi_1, \xi_2, \cdots, \xi_n)'$, $D=[d_1, d_2, \cdots, d_n]$, 则上式化为

$$\xi'D\xi = \sum_{i=1}^{n} d_i\xi_i^2 = 0.$$

据假设, $d_i>0$, $i=1, 2, \cdots, n$, 而 ξ_i 为实数, $1\leqslant i\leqslant n$, 故必 $\xi_1=\cdots=\xi_n=0$, 即 $\xi=0$, 此为矛盾。所以 $|A+D|\neq 0$。

再证 $|A+D|>0$, 作实闭区间$[0, 1]$上的连续实值函数: $f(x)=|xA+D|$, 由于对任意实数 x, xA 仍是反对称阵,故对任何 $x\in[0, 1]$, 恒有 $f(x)=|xA+D|\neq 0$, 由于

$$f(0)=|D|=d_1d_2\cdots d_n>0,$$

故若 $f(1)=|A+D|<0$, 则由连续函数的性质, 必存在 ξ, $0<\xi<1$, 使 $f(\xi)=|\xi A+D|=0$, 此为矛盾,故 $f(1)=|A+D|$ 必大于零。 证毕。

三、用消去法解 $m\times n$ 方程组

本段介绍的方法既可以判断 $m\times n$ 方程 $Ax=b$ 是否有解，且在 $Ax=b$ 相容时，又可将它的全部解找出，这比定理 2.2 提供的解法更实用。先证：

引理 设 A 是秩为 r 的 $m\times n$ 阵，则必存在 m 阶非异阵 P，n 阶排列阵 T，使

$$PAT=\begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix}. \tag{6}$$

【证明】 把定理 1.7 改写成

$$P_1A=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1}=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Q_1 & Q_2 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} Q_1, & Q_2 \\ 0 & 0 \end{pmatrix}, \tag{7}$$

其中 P_1、Q 分别是 m 阶、n 阶非异阵。

由于 A 的秩为 r，故由定理 1.6 的推论及(7)式知，

$$\begin{pmatrix} Q_1 & Q_2 \\ 0 & 0 \end{pmatrix}$$

的秩为 r，设它的 $j_1, j_2, \cdots, j_r$ 个列线性无关，于是存在 n 阶排列阵 $P_{j_1j_2\cdots j_rj_{r+1}\cdots j_n}=T$，使

$$P_1AT=\begin{pmatrix} Q_1 & Q_2 \\ 0 & 0 \end{pmatrix}T=\begin{pmatrix} A_1 & A_2 \\ 0 & 0 \end{pmatrix}, \tag{8}$$

其中 A_1 是 r 阶非异阵，记 $P=[A_1^{-1}, I_{n-r}]P_1$，$A_1^{-1}A_2=B$，于是由(8)式即得

$$PAT=\begin{pmatrix} A_1^{-1} & 0 \\ 0 & I_{n-r} \end{pmatrix}P_1AT=\begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix}, \qquad \text{证毕.}$$

由上述引理，可把 $m\times n$ 方程组 $Ax=b$ 化成同解方程组

$$(PAT)(T'x)=Pb \tag{9}$$

(因 T 是正交阵)，记

$$T'x=Z=\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \quad Pb=C=\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}, \tag{10}$$

由(6)式与(10)式，(9)式可化为同解方程组：

$$\begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}=\begin{pmatrix} C_1 \\ C_2 \end{pmatrix},$$

也即

$$\begin{pmatrix} Z_1+BZ_2 \\ 0 \end{pmatrix}=\begin{pmatrix} C_1 \\ C_2 \end{pmatrix},$$

由此可知，$Ax=b$ 相容的充要条件是 $C_2=0$，当 $C_2=0$ 时，方程组 $Ax=b$ 的解可由解

$$Z_1=C_1-BZ_2 \tag{11}$$

$$x=TZ \tag{12}$$

而得出，此处的 Z_2 可以取任意的 $n-r$ 维列向量，当 $r=n$ 时，则 $Ax=b$ 有唯一解；当 $r<n$ 时 $Ax=b$ 有无穷多个解。

解一个具体的 $m\times n$ 方程组时，根本不必求出 P 与 T 的具体形状，只要用消去法（初等变换）把 A 的增广矩阵 $\tilde{A}=(A,\ b)$ 化成：

$$\left(\begin{array}{cc|c} I_r & B & C_1 \\ 0 & 0 & C_2 \end{array}\right),$$

则当 $C_2=0$ 时，就可由(11)、(12)式找出 $Ax=b$ 的全部解，不过要注意，虽然对 $(A,\ b)$ 的行可以作任何一种初等变换，但对 $(A,\ b)$ 的列，仅能作 A 的列的对调。（当 $C_2\neq 0$ 时，$Ax=b$ 无解。）

例 5 解 3×4 方程组

$$\begin{cases} -4x_1+2x_2-2x_3+\ \ x_4=2 \\ \ \ 2x_1-\ \ x_2+\ \ x_3+3x_4=-8 \\ \ \ 6x_1-3x_2+3x_3+2x_4=-9 \end{cases}$$

【解】 作初等变换

$$\overset{(1)}{}\left(\begin{array}{cccc|c} -4 & 2 & -2 & 1 & 2 \\ 2 & -1 & 1 & 3 & -8 \\ 6 & -3 & 3 & 2 & -9 \end{array}\right)$$

（第1行乘 (-1) 加到第3行）

$$\to\left(\begin{array}{cccc|c} -4 & 2 & -2 & 1 & 2 \\ 2 & -1 & 1 & 3 & -8 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right),$$

由此可知，原方程组必无解。

例 6 解 3×4 方程组

$$\begin{pmatrix} -4 & 2 & -2 & 1 \\ 2 & -1 & 1 & 3 \\ 6 & -3 & 3 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=\begin{pmatrix} 2 \\ -8 \\ -10 \end{pmatrix}.$$

【解】 作增广矩阵的初等变换：

$$\left(\begin{array}{cccc|c} -4 & 2 & -2 & 1 & 2 \\ 2 & -1 & 1 & 3 & -8 \\ 6 & -3 & 3 & 2 & -10 \end{array}\right) \to \left(\begin{array}{cccc|c} -4 & 2 & -2 & 1 & 2 \\ 2 & -1 & 1 & 3 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to$$

$$\to \left(\begin{array}{cccc|c} 1 & 2 & -2 & -4 & 2 \\ 3 & -1 & 1 & 2 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(-\frac{1}{7}\right) \to$$

$$\to \left(\begin{array}{cccc|c} 1 & 2 & -2 & -4 & 2 \\ 0 & -7 & 7 & 14 & -14 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to$$

$$\to (-2)\left(\begin{array}{cccc|c} 1 & 2 & -2 & -4 & 2 \\ 0 & 1 & -1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -2 \\ 0 & 1 & -1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

故原方程组有解，且可化为：

$$z_1 = -2,$$
$$z_2 = 2 + z_3 + 2z_4,$$

而 z_3 与 z_4 可以取任意数，由于

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = P_{14}\begin{pmatrix} 1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} = \begin{pmatrix} z_4 \\ z_2 \\ z_3 \\ z_1 \end{pmatrix} \qquad (\text{此处 } T = P_{4231} = P_{14}),$$

所以原方程组的解是：$x_1 = z_4$，$x_2 = z_2$，$x_3 = z_3$，$x_4 = z_1$，即

$$x_4 = -2,\ x_2 = 2 + x_3 + 2x_1,$$

而 x_1、x_3 可以取任意数。

习　题

1. 判断下列各方程组是否有解，若有解，求出其解。

(i) $\begin{cases} x_1+x_2-x_3=1 \\ 2x_1-3x_2+4x_3=3 \\ x_1-4x_2+5x_3=2 \\ 3x_1-2x_2+3x_3=4; \end{cases}$　(ii) $\begin{cases} x_2+x_3-x_4=0 \\ 2x_1+x_2+3x_3+x_4=1 \\ x_1+x_2+x_3 \qquad =\dfrac{1}{2}; \end{cases}$

(iii) $\begin{cases} 2x_1+x_2-x_3+x_4=1 \\ 3x_1-2x_2+2x_3-3x_4=2 \\ 5x_1+x_2-x_3+2x_4=-1 \\ 2x_1-x_2+x_3-3x_4=4; \end{cases}$　(iv) $\begin{cases} x_1+x_2 = 0 \\ x_2+x_3 = 0 \\ \cdots\cdots \\ x_{n-1}+x_n=0. \end{cases}$

2. 找出线性方程组：

$$x_1-x_2=a_1,\ x_2-x_3=a_2,\ \cdots,\ x_{n-1}-x_n=a_{n-1},\ x_n-x_1=a_n$$

有解的充要条件，並在相容的情形下，求出它的全部解。

3. 设A是n阶阵，b是n维列向量，且

$$r\begin{pmatrix} A & b \\ b' & 0 \end{pmatrix}=r(A),$$

求证：$Ax=b$必有解。反之是否正确？若不正确，试举出反例。

4. (i) 设S是n阶反对称阵，I_n是n阶单位阵，求证：$(I_n-S)(I_n+S)^{-1}$是正交阵。

(ii) 设A是n阶(实)对称阵，求证：$I_n-\sqrt{-1}A$与$I_n+\sqrt{-1}A$都是非异阵，且$(I_n-\sqrt{-1}A)(I_n+\sqrt{-1}A)^{-1}$是酉阵。

5 (i) 如果n阶正交阵A使$A+I_n$是非异阵，求证：必存在n阶反对称阵S，使$A=(I_n-S)(I_n+S)^{-1}$；

(ii) (广义 Cayley 分解)求证：任何n阶正交阵A必有分解式：$A=D(I_n-S)(I_n+S)^{-1}$，其中S是反对称阵，D是对角阵，它的主对角元是1或-1。

(提示：应用§1的习题第11题以及本题(i))

6. 设B是非异对称阵，C是反对称阵，且$BC=CB$，求证：$B+C$是非异阵。当$|B|>0$，则$|B+C|>0$。

7. 设$A=(a_{ij})_{n\times n}$是实方阵，且

$$a_{ii}>\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|,\ i=1,2,\cdots,n,$$

求证：$|A|>0$.

(提示：作[0, 1]上的连续实变实值函数：

$$f(x)=\begin{vmatrix} a_{11} & xa_{12}\cdots xa_{1n} \\ xa_{21} & a_{22} \quad \cdots xa_{2n} \\ \cdots\cdots & \cdots\cdots\cdots \\ xa_{n1} & xa_{n2}\cdots a_{nn} \end{vmatrix},$$

然后应用 Lévy-Hadamard 定理並仿例 4 的做法.)

8. (i) 设A是n阶实方阵，求证：必存在正整数N，使$|NI_n+A|>0$；

(ii) 设A是n阶实方阵，若对任意n维实的列向量x恒有：$x'Ax\neq0$，求证：$|A|\neq0$。

(iii) 设A是n阶实方阵，若对任意n维实的列向量x恒有：$x'Ax>0$，求证：$|A|>0$。

(提示：由 Lévy-Hadamard 定理或直接用第 7 题即得 (i)。又用反证法即得(ii)。对于(iii)，作 $[0, \infty)$ 上的连续实值函数：$f(t)=|tI_n+A|$，余下仿例 4 的证法.)

9. 设A是n阶非异实方阵，S是n阶反对称阵，求证：$|AA'\pm S|>0$.

10. 用本节第三段的引理证明定理 2.1.

(提示：由引理，存在P与T，使

$$PAT=\begin{pmatrix} I_r & B \\ 0 & 0 \end{pmatrix},$$

故

$$r(A, b)=r\left(P(A, b)\begin{pmatrix} T & 0 \\ 0 & 1 \end{pmatrix}\right)=r\ (PAT,\ Pb)\ =r\begin{pmatrix} I_r & B & C_1 \\ 0 & 0 & C_2 \end{pmatrix},$$

于是 $Ax=b$ 相容 $\Longleftrightarrow C_2=0 \Longleftrightarrow r(A, b)=r=r(A)$.)

§3 线性代数方程组的解的结构

定义 称 $m\times n$ 方程组 $Ax=b(b\neq0)$为**非齐次(线性)方程组**，而称 $Ax=0$ 为**齐次(线性)方程组。**

对某一个 $m\times n$ 阵 A，称 $Ax=0$ 为 $Ax=b$ 的**导出方程组。**

本节恒假定 $Ax=b$ 有解。

所谓方程组的解的结构，首先要讨论 $Ax=b$ 的解与它的导出方程组的解的关系，对有唯一解的 $m\times n$ 方程组 $Ax=b$，这个问题

的意义不大，因为此时A的秩$r=n$，因而$Ax=0$只有零解，故仅当$r<n$时，讨论这一问题才有意义。其次，当$r<n$时，我们希望$Ax=b$、特别是$Ax=0$的无穷多个解能够用它的有限个解的某个一般表达式表示出来，找出这种表达式就是解的结构要讨论的第二个问题。

现在先讨论第二个问题，有下面的：

定理 3.1 设$m\times n$方程组$Ax=0$的系数矩阵A的秩为r，$r<n$，则$Ax=0$必有$n-r$个线性无关的解（向量）$\beta_1,\beta_2,\cdots,\beta_{n-r}$，使$Ax=0$的任一解都是$\beta_1,\beta_2,\cdots,\beta_{n-r}$的线性组合，並且，当$k_1,k_2,\cdots,k_{n-r}$遍取任何数时，则$k_1\beta_1+k_2\beta_2+\cdots+k_{n-r}\beta_{n-r}$就是$Ax=0$的全部解。

【证明】 设$A=(a_{ij})_{m\times n}$的第i_1、i_2、…、i_r行构成极大线性无关组，且

$$\left|A\begin{pmatrix}i_1 & i_2\cdots i_r\\ j_1 & j_2\cdots j_r\end{pmatrix}\right|\neq 0,$$

则必存在m阶非异阵P，n阶排列阵$T=P_{j_1j_2\cdots j_rj_{r+1}\cdots j_n}$，使

$$PAT=\begin{pmatrix}A_r & B\\ 0 & 0\end{pmatrix} \tag{1}$$

其中，

$$A_r=A\begin{pmatrix}i_1 & i_2\cdots i_r\\ j_1 & j_2\cdots j_r\end{pmatrix},\quad B=\begin{pmatrix}a_{i_1j_{r+1}}\cdots a_{i_1j_n}\\ \cdots\cdots\cdots\cdots\\ a_{i_rj_{r+1}}\cdots a_{i_rj_n}\end{pmatrix},$$

于是$Ax=0$可化为同解方程组：$(PAT)(T'x)=0$，即

$$\begin{pmatrix}A_r & B\\ 0 & 0\end{pmatrix}\begin{pmatrix}x_{j_1}\\ \vdots\\ x_{j_r}\\ x_{j_{r+1}}\\ \vdots\\ x_{j_n}\end{pmatrix}=A_r\begin{pmatrix}x_{j_1}\\ x_{j_2}\\ \vdots\\ x_{j_r}\end{pmatrix}+B\begin{pmatrix}x_{j_{r+1}}\\ x_{j_{r+2}}\\ \vdots\\ x_{j_n}\end{pmatrix}=0 \tag{2}$$

或者，

$$\begin{cases} a_{i_1j_1}x_{j_1}+a_{i_1j_2}x_{j_2}+\cdots+a_{i_1j_r}x_{j_r}+a_{i_1j_{r+1}}x_{j_{r+1}}+\cdots+a_{i_1j_n}x_{j_n}=0 \\ a_{i_2j_1}x_{j}+a_{i_2j_2}x_{j_2}+\cdots+a_{i_2j_r}x_{j_r}+a_{i_2j_{r+1}}x_{j_{r+1}}+\cdots+a_{i_2j_n}x_{j_n}=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i_rj_1}x_{j_1}+a_{i_rj_2}x_{j_2}+\cdots+a_{i_rj_r}x_{j_r}+a_{i_rj_{r+1}}x_{j_{r+1}}+\cdots+a_{i_rj_n}x_{j_n}=0 \end{cases}$$

由于 A_r 是非异阵，故由(2)即得

$$\begin{pmatrix} x_{j_1} \\ x_{j_2} \\ \vdots \\ x_{j_r} \end{pmatrix}=(-A_r^{-1}B)\begin{pmatrix} x_{j_{r+1}} \\ x_{j_{r+1}} \\ \vdots \\ x_{j_n} \end{pmatrix},$$

于是

$$x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=T\begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_r} \\ x_{j_{r+1}} \\ \vdots \\ x_{j_n} \end{pmatrix}=T\begin{pmatrix} -A_r^{-1}B \\ I_{n-r} \end{pmatrix}\begin{pmatrix} x_{j_{r+1}} \\ x_{j_{r+2}} \\ \vdots \\ x_{j_n} \end{pmatrix}, \tag{3}$$

这说明，$Ax=0$ 的解 x 应是(3)的形状，又由于(1)式，故

$$AT\begin{pmatrix} -A_r^{-1}B \\ I_{n-r} \end{pmatrix}\begin{pmatrix} x_{j_{r+1}} \\ x_{j_{r+2}} \\ \vdots \\ x_{j_n} \end{pmatrix}=P^{-1}\begin{pmatrix} A_r & B \\ 0 & 0 \end{pmatrix}\begin{pmatrix} -A_r^{-1}B \\ I_{n-r} \end{pmatrix}\begin{pmatrix} x_{j_{r+1}} \\ x_{j_{r+2}} \\ \vdots \\ x_{j_n} \end{pmatrix}$$

$$=P^{-1}\begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} x_{j_{r+1}} \\ x_{j_{r+2}} \\ \vdots \\ x_{j_n} \end{pmatrix}=0,$$

所以对任何数 $x_{j_{r+1}}, x_{j_{r+2}}, \cdots, x_{j_n}$，(3) 的右边的 n 维向量就是 $Ax=0$ 的解。

今把$\begin{pmatrix} -A_r^{-1}B \\ I_{n-r} \end{pmatrix}$按它的列分块，$\begin{pmatrix} -A_r^{-1}B \\ I_{n-r} \end{pmatrix}=(\delta_1, \delta_2, \cdots, \delta_{n-r})$，

则 $\delta_1, \delta_2, \cdots, \delta_{n-r}$ 线性无关(由定理 1.1)，记

$$\beta_j=T\delta_j,\ j=1, 2, \cdots, n-r,$$

则 $\beta_1, \beta_2, \cdots, \beta_{n-r}$ 也线性无关。又因(3)的右边的 $(x_{j_{r+1}}, x_{j_{r+2}},$

$\cdots, x_{j_n})'$ 可以取任何列向量，故分别取 $(x_{j_{r+1}}, x_{j_{r+2}}, \cdots, x_{j_n})' = e_1, e_2, \cdots, e_{n-r}$，则可知 $\beta_1, \beta_2, \cdots, \beta_{n-r}$ 全是 $Ax=0$ 的解。

又若 $\alpha=(a_1, \cdots, a_r, a_{r+1}, \cdots, a_n)'=T(a_{j_1}, \cdots, a_{j_r}, a_{j_{r+1}}, \cdots, a_{j_n})'$ 是 $Ax=0$ 的任一解，则由(3)可得

$$\alpha = T\begin{pmatrix} -A_r^{-1}B \\ I_{n-r} \end{pmatrix}\begin{pmatrix} a_{j_{r+1}} \\ a_{j_{r+2}} \\ \vdots \\ a_{j_n} \end{pmatrix} = (\beta_1, \beta_2, \cdots, \beta_{n-r})\begin{pmatrix} a_{j_{r+1}} \\ a_{j_{r+2}} \\ \vdots \\ a_{j_n} \end{pmatrix}$$

$$= a_{j_{r+1}}\beta_1 + a_{j_{r+2}}\beta_2 + \cdots + a_{j_n}\beta_{n-r},$$

故 α 是 $\beta_1, \beta_2, \cdots, \beta_{n-r}$ 的线性组合。

另一方面，由于 $A\beta_i=0$，$i=1, 2, \cdots, n-r$，故对任何数 $k_1, k_2, \cdots, k_{n-r}$，显然成立：

$$A\left(\sum_{j=1}^{n-r} k_j\beta_j\right) = \sum_{j=1}^{n-r} k_j(A\beta_j) = 0^{①},$$

所以当 $k_1, k_2, \cdots, k_{n-r}$ 取遍所有数时，$\sum_{j=1}^{n-r} k_j\beta_j$ 就是 $Ax=0$ 的全部解。 证毕。

由定理 3.1，我们引进

定义 齐次线性方程组 $Ax=0$ 的一组解 $\beta_1, \beta_2, \cdots, \beta_p$ 称为 $Ax=0$ 的一个**基础解系**，如果

(i) $\beta_1, \beta_2, \cdots, \beta_p$ 线性无关；

(ii) $Ax=0$ 的任一解都可表为 $\beta_1, \beta_2, \cdots, \beta_p$ 的线性组合。

由定理 3.1 可知，$Ax=0$ 的基础解系必定存在，由于 $Ax=0$ 的极大独立(无关)方程组可能不止一组，所以 $Ax=0$ 的基础解系可能不止一个，但由定理 1.2 的推论 3 可知，$Ax=0$ 的任何两个基础解系所含的向量个数必为同一个 p。当 A 的秩为 r 时，由定理 1.3 可知，$p=n-r$。

例 1 由于方程组

---

① 记号 $\sum_{j=1}^{s}\alpha_j$ 是关于 s 个向量 α_j 求和，其定义与性质同第一章 §1。

$$\begin{cases} -2x_1-3x_2-x_3+9x_4=0 \\ 3x_1+2x_2+x_3-4x_4=0 \\ 4x_1+\ \ x_2+x_3+\ \ x_4=0 \end{cases}$$

的系数矩阵的秩为 2，且

$$\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \neq 0,$$

故可取它的一个同解方程组：

$$\begin{cases} 2x_2+x_3+3x_1-4x_4=0 \\ \ \ x_2+x_3+4x_1+\ \ x_4=0 \end{cases}$$

上方程组的解为：

$$\begin{pmatrix} x_2 \\ x_3 \\ x_1 \\ x_4 \end{pmatrix}=\begin{pmatrix} -\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 3 & -4 \\ 4 & 1 \end{pmatrix} \\ I_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_4 \end{pmatrix}=\begin{pmatrix} 1 & 5 \\ -5 & -6 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_4 \end{pmatrix},$$

故原方程组的解为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}=P_{2314}\begin{pmatrix} x_2 \\ x_3 \\ x_1 \\ x_4 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 1 & 5 \\ -5 & -6 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_4 \end{pmatrix},$$

其中 x_1, x_4 可取任意数，易知原方程组的基础解系是，

$$(1, 1, -5, 0)', (0, 5, -6, 1)'.$$

例 2 求证：对秩为 $r<n$ 的 $m\times n$ 阵 A，必存在秩为 $n-r$ 的 $n\times(n-r)$ 阵 C，使 $AC=0$。

【证明】 设 $\beta_1, \beta_2, \cdots, \beta_{n-r}$ 是 $Ax=0$ 的一个基础解系，则由 $A(\beta_1, \beta_2, \cdots, \beta_{n-r})=0$ 可知，$C=(\beta_1, \beta_2, \cdots, \beta_{n-r})$ 即为所求。

证毕。

定理 3.2 设 $m\times n$ 方程组 $Ax=b$ 有无穷多个解，α_0 是其中的某一个(解)，又设 $\beta_1, \beta_2, \cdots, \beta_{n-r}$ 是导出方程组 $Ax=0$ 的基础解系(由此显然可知 A 的秩为 r)，则当 $k_1, k_2, \cdots, k_{n-r}$ 遍取任何数

时,向量组

$$\alpha_0+k_1\beta_1+k_2\beta_2+\cdots+k_{n-r}\beta_{n-r} \tag{4}$$

就是 $Ax=b$ 的全部解。

【证明】 设 α 是 $Ax=b$ 的任一解,则 $A\alpha=b$,又由假设,$A\alpha_0=b$,故 $A(\alpha-\alpha_0)=0$,这说明 $\alpha-\alpha_0$ 是 $Ax=0$ 的解,故由定理 3.1,$\alpha-\alpha_0=l_1\beta_1+l_2\beta_2+\cdots+l_{n-r}\beta_{n-r}$,于是 $\alpha=\alpha_0+l_1\beta_1+l_2\beta_2+\cdots+l_{n-r}\beta_{n-r}$,这说明 $Ax=b$ 的任一解都可写成(4)的形状。

反之,对任何数 $k_1, k_2, \cdots, k_{n-r}$,由于

$$A(\alpha_0+k_1\beta_1+k_2\beta_2+\cdots+k_{n-r}\beta_{n-r})$$

$$=A\alpha_0+\sum_{j=1}^{n-r}k_j(A\beta_j)=b+0=b,$$

故综合以上两点可知,(4)就是 $Ax=b$ 的全部解。 证毕。

对任何(参)变数 $k_1, k_2, \cdots, k_{n-r}$,常称 $Ax=b$ 的解的一般表达式(4)为 $Ax=b$ 的**通解**或**一般解**。

如果称 $Ax=b$ 的某一个解为 $Ax=b$ 的**特解**,则定理 3.2 可改述为

定理 3.2′ 若 $Ax=b$ 有无穷多个解,则 $Ax=b$ 的通解等于 $Ax=b$ 的特解加上 $Ax=0$ 的通解

我们常称定理 3.2′ 为解的**第一结构定理**,而称定理 3.1 为解的**第二个结构定理**。

最后举一稍为复杂之例以结束本节

例 3 设 α 与 β 都是 n 维实的列向量,且满足 $\alpha'\beta=1$,试具体找出 n 阶实方阵 A 与 B,它们的第一列分别是 α 与 β,使得 $A'B=I_n$。

【解】 设 $\alpha=(a_1, a_2, \cdots, a_n)'$,$\beta=(b_1, b_2, \cdots, b_n)'$。由 $\alpha'\beta=1$ 可得,

$$b_1a_1+b_2a_2+\cdots+b_na_n=1, \tag{5}$$

故至少存在某个 i,$1\leqslant i\leqslant n$,使 $b_ia_i\neq0$,不失一般性,可设 $b_1a_1\neq0$,故 $b_1\neq0$,于是如下 $1\times n$ 线性方程组:

$$b_1x_1+b_2x_2+\cdots+b_nx_n=0 \tag{6}$$

的基础解系是如下的$n\times(n-1)$阵

$$\begin{pmatrix} -b_1^{-1}b_2 & -b_1^{-1}b_3 & \cdots & -b_1^{-1}b_n \\ 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}=(\alpha_2,\alpha_3,\cdots,\alpha_n) \tag{7}$$

的$n-1$个列向量$\alpha_2,\alpha_3,\cdots,\alpha_n$. 作$n$阶阵:

$$A=(\alpha_1,\alpha_2,\cdots,\alpha_n),\text{ 此处 }\alpha_1=\alpha,$$

由行列式的第一降阶定理易知,

$$|A|=|I_{n-1}|\cdot\left\{a_1+(b_1^{-1}b_2,b_1^{-1}b_3,\cdots,b_1^{-1}b_n)\begin{pmatrix}a_2\\a_3\\\vdots\\a_n\end{pmatrix}\right\}$$

$$=a_1+b_1^{-1}(b_2a_2+b_3a_3+\cdots+b_na_n)=b_1^{-1}\neq0$$

(用了(5)式), 故A是非异阵. 且由(6)与(7)式可知, $\beta'\alpha_i=0$, 即$\alpha_i'\beta=0$, $i=2,3,\cdots,n$. 于是

$$A'\beta=\begin{pmatrix}\alpha_1'\\\alpha_2'\\\vdots\\\alpha_n'\end{pmatrix}\beta=\begin{pmatrix}\alpha_1'\beta\\\alpha_2'\beta\\\vdots\\\alpha_n'\beta\end{pmatrix}=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}=e_1. \tag{8}$$

作n阶阵:

$$B=(\beta,(A')^{-1}e_2,(A')^{-1}e_3,\cdots,(A')^{-1}e_n),$$

则由(8)式知, $A'B=(A'B,e_2,e_3,\cdots,e_n)=I_n$ 故上述A与B即为所求之阵.

习　　题

1. 求下列齐次线性方程组的基础解系:

(i) $$\begin{cases}x_1+x_2-3x_4-x_5=0\\x_1-x_2+2x_3-x_4=0\\4x_1-2x_2+6x_3+3x_4-4x_5=0\\2x_1+4x_2-2x_3+4x_4-7x_5=0;\end{cases}$$

(ii) $\begin{cases} x_1 - x_2 + x_3 - x_4 = 0 \\ 2x_1 - 2x_2 + 2x_3 - x_4 = 0 \\ 3x_1 - 3x_2 + 4x_3 - x_4 = 0 \\ x_1 - x_2 - 2x_3 + x_4 = 0. \end{cases}$

2. 设A是秩为 r 的 $m\times n$ 阵，$r<n$，求证：$Ax=0$ 的任意 $n-r$ 个线性无关的解向量构成 $Ax=0$ 的一个基础解系。

3. 求证：任何 r 个线性无关的 n 维列向量 ($r<n$) 必可作为某个 $m\times n$ 方程组 $Ax=0$ 的基础解系，且 $r(A)=n-r$。

4. 设 $\delta_1, \delta_2, \cdots, \delta_t$ 是 $Ax=0$ 的线性无关解向量，求证：A 的秩不大于 A的列数减去 t。

5. 设 $D=[\lambda I_r, \mu I_{n-r}]$，$\lambda\neq\mu$，求证：$(\lambda I_n - D)x=0$ 的基础解系所含的向量个数为 r。

6. 设 A_1 是秩为 r 的 $r\times n$ 阵A的前 p 行所成的 $p\times n$ 阵，$p<r$，求证：在 $A_1x=0$ 的所有解中，至少存在一个解 δ，使得 $A\delta\neq0$。

7. 设 A_{ij} 是奇异阵 $A=(a_{ij})_{n\times n}$ 的某个元素 a_{ij} 的代数余子式，且 $A_{ij}\neq0$，求证：$Ax=0$ 的任一解必可表为：$(kA_{i1}, \cdots, kA_{in})'$，此处 k 是任一数。

§4 矩阵的秩的理论及应用

一、列满秩阵与行满秩阵

定义 秩为 r 的 $n\times r$ 阵称为**列满秩阵**，秩为 r 的 $r\times n$ 阵称为**行满秩阵**。

满秩阵(非异阵)既是列满秩阵，又是行满秩阵。

定理 4.1 设 r 个 n 维列向量 $\beta_1, \beta_2, \cdots, \beta_r$ 线性无关，则必可找到$n-r$ 个线性无关的列向量 $\beta_{r+1}, \cdots, \beta_n$，使 $\beta_1, \cdots, \beta_r, \beta_{r+1}, \cdots, \beta_n$ 线性无关。换句话说，对任一 $n\times r$ 列满秩阵 $B=(\beta_1, \beta_2, \cdots, \beta_r)$，必可找到$n\times(n-r)$列满秩阵 $C=(\beta_{r+1}, \cdots, \beta_n)$，使 $A=(B, C)$为 n 阶满秩阵。

【证明】 设B的 r 阶子式 $\left|B\begin{pmatrix} i_1 & i_2\cdots i_r \\ 1 & 2\cdots r \end{pmatrix}\right|$ 不等于零，则存在排列阵 $T=p'_{i_1i_2\cdots i_ri_{r+1}\cdots i_n}$，使得

$$TB=\begin{pmatrix} B\begin{pmatrix} i_1 & i_2\cdots i_r \\ 1 & 2\cdots r\end{pmatrix} \\ B_2 \end{pmatrix},$$

由于

$$G=\begin{pmatrix} B\begin{pmatrix} i_1 & i_2\cdots i_r \\ 1 & 2\cdots r\end{pmatrix} & 0 \\ B_2 & I_{n-r} \end{pmatrix}$$

是非异阵，T 是正交阵，故作 $n\times(n-r)$ 阵：

$$C=T'\begin{pmatrix} 0 \\ I_{n-r} \end{pmatrix},$$

易知 C 是列满列阵，且

$$A=(B,\ C)=T'\left(TB,\ \begin{pmatrix} 0 \\ I_{n-r} \end{pmatrix}\right)=T'G$$

是非异阵。　　　　　　　　　　　　　　　　　　　证毕。

定理 4.1 的证明是构造性的，由 B 可以具体找出 C，且从证明过程中可以看出，这种 C 不是唯一的。

定理 4.2　对 $m\times n$ 列满秩阵 B，必存在 $n\times m$ 行满秩阵 A，使 $AB=I_n$。而 $A=(\bar{B}'B)^{-1}\bar{B}'$ 便是其中之一。

【证明】　设 $|B_1|=\left|B\begin{pmatrix} i_1 & i_2\cdots i_n \\ 1 & 2\cdots n\end{pmatrix}\right|\neq 0$，则必存在 m 阶排列阵 $T=P'_{i_1\cdots i_n\cdots i_m}$，使

$$TB=\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}.$$

由于 B_1 是 n 阶非异阵，故记 $A_1=B_1^{-1}$，则

$$(A_1,\ 0)TB=(A_1,\ 0)\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}=A_1B_1=I_n.$$

记 $A=(A_1,\ 0)T$，则 A 显然是 $n\times m$ 行满秩阵，故 A 即为所求之阵。　　　　　　　　　　　　　　　　　　　　　证毕。

对行满秩阵 R，也有与定理 4.2 相平行的结论，即存在列满秩阵 C，使 $RC=I$。

由定理 4.2 及上述说明显然可得

推论

(i) 设 A 是 $m\times n$ 列满秩阵，B 与 C 都是 $n\times l$ 阵，如果 $AB=AC$，则必 $B=C$。换言之，任一列满秩阵可以从矩阵等式的左侧消去；

(ii) 行满秩阵可以从矩阵等式的右侧消去。

由于列满秩阵与行满秩阵是比满秩阵（非异阵）更广泛的概念，所以可用它们处理某些长方矩阵的问题，在本节中将逐步看到这一点，今先举例说明之。

例 1 设 A 是秩为 r 的 n 阶阵，则必存在 n 阶非异阵 P，使

$$P^{-1}AP=\begin{pmatrix}B\\0\end{pmatrix}, \tag{1}$$

其中 B 是 $r\times n$ 行满秩阵。

【证明】 由定理 1.7，存在 n 阶非异阵 R 与 Q，使

$$RAQ=\begin{pmatrix}I_r & 0\\0 & 0\end{pmatrix},$$

记 $R^{-1}=P$，$Q^{-1}P=\begin{matrix} \\ r\\ \\ \end{matrix}\!\!\begin{pmatrix}n\\B\\C\end{pmatrix}$，则上式化为

$$P^{-1}AP=\begin{pmatrix}I_r & 0\\0 & 0\end{pmatrix}\begin{pmatrix}B\\C\end{pmatrix}=\begin{pmatrix}B\\0\end{pmatrix},$$

易知 B 是 $r\times n$ 行满秩阵。 证毕。

例 2 设 A 是秩为 r 的幂等阵（即 $A^2=A$），则必存在非异阵 P，使

$$P^{-1}AP=\begin{pmatrix}I_r & 0\\0 & 0\end{pmatrix}. \tag{2}$$

【证明】 由例 1，存在非异阵 R，使

$$R^{-1}AR=\begin{matrix} \\ r\\ \\ \end{matrix}\!\!\begin{pmatrix}\overset{r}{B_1} & B_2\\0 & 0\end{pmatrix}. \tag{3}$$

由假设，$A^2=A$，故得

$$\begin{pmatrix} B_1 & B_2 \\ 0 & 0 \end{pmatrix} = R^{-1}AR = (R^{-1}AR)^2$$

$$= \begin{pmatrix} B_1 & B_2 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} B_1^2 & B_1B_2 \\ 0 & 0 \end{pmatrix},$$

于是

$$(B_1, B_2) = (B_1^2, B_1B_2) = B_1(B_1, B_2),$$

也即

$$(B_1 - I_r)(B_1, B_2) = 0 = 0(B_1, B_2),$$

由于(B_1, B_2)是行满秩阵，故可从上式的右侧消去，于是$B_1 = I_r$，代入(3)式即得

$$R^{-1}AR = \begin{pmatrix} I_r & B_2 \\ 0 & 0 \end{pmatrix}, \tag{4}$$

把等式(4)右边的矩阵作分块阵的初等变换及其相应的逆变换，即

$$T_{12}(B_2)\begin{pmatrix} I_r & B_2 \\ 0 & 0 \end{pmatrix} T_{12}(-B_2) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

记 $P = RT_{12}(-B_2)$，则由(4)式与上式即得(2)式。　　　　证毕。

由(2)式可推出幂等阵的一些性质，例如：

(i) $|I_n - 2A| = (-1)^r$（即由 r 的奇偶性定出对合阵 $I_n - 2A$ 的行列式的正负号）。

(ii) $\mathrm{r}(A) + \mathrm{r}(I_n - A) = n$。　　(5)

(iii) 任何实幂等阵必可分解为两个(实)对称阵的乘积。

(i)与(ii)容易证得。今证(iii)：因(2)式可写成，

$$A = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P^{-1} = \left(P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}P'\right)\left((P^{-1})'P^{-1}\right) = S_1S_2,$$

其中 $S_1 = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}P'$，$S_2 = (P^{-1})'P$，易知它们都是(实)对称阵，且由于 P 可具体找出，故 S_1 与 S_2 都可具体构造出来。

二、矩阵的秩的降阶定理

定理 4.3(第一降阶定理) 设A是非异阵，$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$是$m\times n$阵，则

$$\mathrm{r}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mathrm{r}(A) + r(D - CA^{-1}B). \tag{6}$$

【证明】 因为

$$T_{21}(-CA^{-1})\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix},$$

而A是非异阵，故

$$\begin{aligned}
\mathrm{r}\begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \mathrm{r}\left(T_{21}(-CA^{-1})\begin{pmatrix} A & B \\ C & D \end{pmatrix}\right) \\
&= \mathrm{r}\begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix} \\
&= \mathrm{r}(A) + \mathrm{r}(D - CA^{-1}B),
\end{aligned}$$

证毕。

显然可把(6)式改写成

$$\mathrm{r}(D - CA^{-1}B) = \mathrm{r}\begin{pmatrix} A & B \\ C & D \end{pmatrix} - \mathrm{r}(A), \tag{7}$$

称(7)式为**秩的升阶公式**。

推论 设A与B分别是$m\times n$阵与$n\times l$阵，则

$$\mathrm{r}(AB) \geqslant \mathrm{r}(A) + \mathrm{r}(B) - n. \tag{8}$$

【证明】 由秩的升阶公式，

$$\begin{aligned}
\mathrm{r}(AB) = \mathrm{r}(0 + AI_n^{-1}B) &= \mathrm{r}\begin{pmatrix} I_n & B \\ -A & 0 \end{pmatrix} - r(I_n) \\
&= \mathrm{r}\begin{pmatrix} B & -I_n \\ 0 & A \end{pmatrix} - r(I_n) \geqslant \mathrm{r}(A) + \mathrm{r}(B) - n.
\end{aligned}$$

证毕。

把定理 1.6 与(8)式合并，即得矩阵乘积的秩的上、下界估计式：

$$\min(\mathrm{r}(A), \mathrm{r}(B)) \geqslant \mathrm{r}(AB) \geqslant \mathrm{r}(A) + \mathrm{r}(B) - A \text{ 的列数}, \tag{9}$$

常称(9)式为 **Sylvester 不等式**，它在矩阵的秩的理论中是个常用的结果。

例 3 设 A 与 B 分别是 $m\times n$ 阵与 $n\times l$ 阵，且 $AB=0$，则 $\mathrm{r}(A)\leqslant n-\mathrm{r}(B)$。

【证明】 由(8)式及假设：$AB=0$ 即得，$0=\mathrm{r}(AB)\geqslant \mathrm{r}(A)+\mathrm{r}(B)-n$，于是 $\mathrm{r}(A)\leqslant n-\mathrm{r}(B)$。 证毕。

例 4 设 A 是 n 阶阵，B 与 C 分别是 $m\times l$ 阵与 $n\times l$ 阵，则

$$\mathrm{r}(A)+\mathrm{r}(C)\leqslant \mathrm{r}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}\leqslant \min(m+\mathrm{r}(C),\ l+\mathrm{r}(A))。\quad (10)$$

【证明】 由 Sylvester 不等式：

$$\min\left(\begin{pmatrix} I_m & 0 \\ 0 & C \end{pmatrix}, \mathrm{r}\begin{pmatrix} A & B \\ 0 & I_l \end{pmatrix}\right)\geqslant \mathrm{r}\left(\begin{pmatrix} I_m & 0 \\ 0 & C \end{pmatrix}\begin{pmatrix} A & B \\ 0 & I_l \end{pmatrix}\right)$$

$$=\mathrm{r}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}, \quad (11)$$

并且 $$\mathrm{r}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}=\mathrm{r}\left(\begin{pmatrix} I_m & 0 \\ 0 & C \end{pmatrix}\begin{pmatrix} A & B \\ 0 & I_l \end{pmatrix}\right)\geqslant \mathrm{r}\begin{pmatrix} I_m & 0 \\ 0 & C \end{pmatrix}$$

$$+\mathrm{r}\begin{pmatrix} A & B \\ 0 & I_l \end{pmatrix}-(m+l),$$

但是

$$\mathrm{r}\begin{pmatrix} I_m & 0 \\ 0 & C \end{pmatrix}=m+\mathrm{r}(C),\ \mathrm{r}\begin{pmatrix} A & B \\ 0 & I_l \end{pmatrix}=\mathrm{r}(A)+l,$$

故得

$$\mathrm{r}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}\geqslant \mathrm{r}(A)+\mathrm{r}(C), \quad (12)$$

联合(11)式与(12)式即得(10)式。 证毕。

由 Sylvester 不等式可得一个有用的推论，它是定理 1.6 的推论。

推论 设 A 是 $m\times n$ 阵，P 是 $l\times m$ 列满秩阵，Q 是 $n\times s$ 行满秩阵，则

$$\mathrm{r}(PA)=\mathrm{r}(AQ)=\mathrm{r}(PAQ)=\mathrm{r}(A)。\quad (13)$$

【证明】 由 Sylvester 不等式，

$$\mathrm{r}(A)\geqslant \mathrm{r}(PA)\geqslant \mathrm{r}(P)+\mathrm{r}(A)-m=\mathrm{r}(A),$$

故 $r(PA)=r(A)$。同理可证：$r(AQ)=r(A)$，于是，

$$r(PAQ)=r(AQ)=r(A)。$$ 证毕。

与定理 4.3 的证明相仿，可得下面的

定理 4.4 设 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 是 $m\times n$ 阵，D 是非异阵，则

$$r\begin{pmatrix} A & B \\ C & D \end{pmatrix}=r(D)+r(A-BD^{-1}C) \tag{14}$$

仍称此定理为秩的第一降阶定理。

当 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 是方阵，且 A 与 D 都是非异阵，则由(6)式与(14)式显然可得

定理 4.5（第二降阶定理） 设 A 与 D 分别是 r 阶与 s 阶非异阵(未必同阶)，B 与 C 分别是 $r\times s$ 阵与 $s\times r$ 阵，则

$$r(D-CA^{-1}B)=r(D)-r(A)+r(A-BC^{-1}C)。 \tag{15}$$

例 5 由第二降阶定理可知，对 $s\times n$ 复矩阵 A 恒有

$$r(I_s-A\overline{A'})=s-n+r(I_n-\overline{A'}A)$$

也即

$$r(I_s-A\overline{A'})+n=r(I_n-\overline{A'}A)+s。$$

用两个降阶定理求数字矩阵的秩也是方便的，就不再举例了。

三、矩阵的满秩分解

定理 4.6 设 A 是秩为 r 的 $m\times n$ 阵，则

(i) $$A=HL, \tag{16}$$

其中 H 是 $m\times r$ 列满秩阵，L 是 $r\times n$ 行满秩阵，称(16)式为 A 的**满秩分解**。

(ii) 如果 $A=HL=H_1L_1$ 是 A 的任意两个不同的满秩分解，则必存在 r 阶非异阵 P，使

$$H=H_1P,\ L=P^{-1}L_1。 \tag{17}$$

【证明】 由定理 1.7，存在 m 阶非异阵 P 与 n 阶非异阵 Q，使

$$A = P^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} \tag{18}$$

把 P^{-1} 与 Q^{-1} 如下分块：

$$P^{-1} = \overset{\quad r \qquad n-r}{(H, \quad H_1)}, \quad Q^{-1} = \begin{matrix} r \\ n-r \end{matrix}\begin{pmatrix} L \\ L_1 \end{pmatrix},$$

代入(18)式即得

$$A = (H, 0)\begin{pmatrix} L \\ L_1 \end{pmatrix} = HL.$$

由于 P^{-1} 与 Q^{-1} 均为非异阵，故 P^{-1} 的前 r 列与 Q^{-1} 的前 r 行显然都分别线性无关，所以 $\mathrm{r}(H) = \mathrm{r}(L) = r$。这证明了(i)。

再证(ii)。对 $r\times n$ 行满秩阵 L，存在 $n\times r$ 列满秩阵 N，使 $LN = I_r$，故得

$$H = HI_r = H(LN) = (HL)N = H_1(L_1N). \tag{19}$$

记 $L_1N = P$，则(19)式化为

$$H = H_1P. \tag{20}$$

由于 H 是 $m\times r$ 列满秩阵，故存在 $r\times n$ 行满秩阵 M，使 $MH = I_r$（由定理 4.2），于是在(20)两边左乘 M，得到

$$I_r = (MH_1)P,$$

上式说明 P 是(r 阶)非异阵，且 $P^{-1} = MH_1$，于是

$$L = (MH)L = (MH_1)L_1 = P^{-1}L_1.$$

证毕。

引理 设 A 是 r 阶非异阵，则 $m\times n$ 阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 的秩为 r 的充要条件是，$D - CA^{-1}B = 0$。

这由秩的第一降阶定理立刻可以看出。

当 $m\times n$ 阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 与非异阵 A 有相同的秩 r 时，则由引理可知

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C & CA^{-1}B \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix}(I_r, A^{-1}B)$$

$$= \begin{pmatrix} I_r \\ CA^{-1} \end{pmatrix}(A, B),$$

故由上式即得$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$的两种满秩分解：

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I_r \\ CA^{-1} \end{pmatrix}(A, B) = H_1 L_1, \tag{21}$$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix}(I_r, A^{-1}B) = HL. \tag{22}$$

易知 $H = \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} I_r \\ CA^{-1} \end{pmatrix} A = H_1 A$，而 $L = (I_r, A^{-1}B) = A^{-1}(A, B)$ $= A^{-1}L_1$。

例 6 下列 3 阶阵的秩与它的 2 阶顺序主子阵的秩均为 2，故由(22)与(21)式，它有满秩分解：

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 2 \\ 0 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{pmatrix}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}^{-1}\begin{pmatrix} -1 \\ 2 \end{pmatrix}\right)$$

$$= \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 4 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 2 \\ 0 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ (0, 1)\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}^{-1} \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 2 \end{pmatrix}.$$

对一般的 $m \times n$ 阵，按照定理 4.6 证明中提供的构造性方法，用初等变换同时求出 $m \times n$ 阵的秩以及相应的 P^{-1} 与 Q^{-1}（参考第一章 §5），便可找出它的满秩分解，这里就不再举具体之例了。

例 7 秩为 1 的方阵 $A = (a_{ij})_{n \times n}$ 必满足：$A^2 = \left(\sum_{i=1}^{n} a_{ii}\right)A$。

这是因为，当 $r(A) = 1$ 时，由定理 4.6，

$$A=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1, b_2, \cdots, b_n),$$

此处$(a_1, a_2, \cdots, a_n)'$与$(b_1, b_2, \cdots, b_n)$均为非零向量，且$a_{ii}=a_ib_i$，$i=1, 2, \cdots, n$。于是

$$A^2=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1, b_2, \cdots, b_n)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1, b_2, \cdots, b_n)$$

$$=\left(\sum_{i=1}^{n}a_ib_i\right)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1, b_2, \cdots, b_n)=\left(\sum_{i=1}^{n}a_{ii}\right)A.$$

推论 设A是$m\times n$复矩阵，则

$$r(A\bar{A}')=r(\bar{A}'A)=r(A). \tag{23}$$

【证明】 设$r(A)=r$，由定理4.6，$A=HL$，其中H与L分别是$m\times r$列满秩阵与$r\times n$行满秩阵。连续运用Sylvester不等式的推论((13式)，得到

$$r(\bar{A}'A)=r(\bar{L}'\bar{H}'HL)=r(\bar{H}'H). \tag{24}$$

因为$r(H)=r$，故H中至少有一个r阶子式不等于零，故由Cauchy-Binet公式，$|\bar{H}'H|>0$，这说明$r(\bar{H}'H)=r=r(A)$，所以由(24)式即得$r(\bar{A}'A)=r(A)$。同理可证：$r(A\bar{A}')=r(A)$。 证毕。

矩阵的满秩分解有较多的应用，将在本章选做题中看到这一点。最后举一个综合运用秩的理论之例。

例8 求证：n阶阵A是幂等阵的充要条件为

$$r(A)+r(I_n-A)=n.$$

【证明】 必要性已见之于例2的(ii)(即(5)式)。今证充分性。设$r(A)=r$，$A=HL$是满秩分解，其中H与L分别是$n\times r$列满秩阵与$r\times n$行满秩阵，于是由秩的第二降阶定理，

$$r(I_n-A)=r(I_n-HL)=n-r+r(I_r-LH),$$

由上式及充分条件的假设可知，

$$r(I_r-LH)=r(I_n-A)+r(A)-n=0,$$

故得 $LH=I_r$，所以

$$A^2=HLHL=HI_rL=HL=A.$$ 证毕。

若把上面的证明倒推回去，则显然又可证得必要性。由此例还看出，若不运用秩的第二降阶定理以及矩阵的满秩分解，充分性的证明将较为麻烦。

习　　题

1. 设 A 是 n 阶对合阵(即 $A^2=I_n$)，且 I_n+A 的秩为 r，求证：

(i) 必存在非异阵 P，使

$$P^{-1}AP=\begin{pmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{pmatrix};$$

(ii) $|I_n+2A|>0$ 的充要条件是，n 与 r 有相同的奇偶性；

(iii) $r(I_n-A)+r(I_n+A)=n$；

(iv) 当 A 是实对合阵时，则 A 必可分解为两个(实)对称阵的乘积。

2. 如果 n 阶阵 A 满足条件：$r(I_n-A)+r(I_n+A)=n$，求证：A 必是对合阵。

3. 设 n 阶阵 A 的秩为 r，且 $A^3=A$，求证：必存在非异阵 P，使

$$P^{-1}AP=\begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_{r-s} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

其中 $s\leqslant r$，当 $s=r$ 时，规定 $I_0=0$。

4. (i) 设 A 是秩为 r 的 n 阶阵，且 $A^2=kA$，$k\neq 0$，求证：必存在非异阵 P，使

$$P^{-1}AP=\begin{pmatrix} kI_r & 0 \\ 0 & 0 \end{pmatrix};$$

(ii) 设 A 与 B 都是 n 阶阵：

$$A=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \quad B=\begin{pmatrix} n & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

求证：必存在 n 阶非异阵 P，使 $P^{-1}AP=B$。

5. 设A是n阶阵，$A^2=kI_n$，$k\neq0$，求证：必存在n阶非异阵P，使$P^{-1}AP$是对角阵，並求出这个对角阵的具体形状

6. 设A是n阶2次幂零阵，即$A^2=0$，求证：必存在非异阵P，使

$$P^{-1}AP=\begin{pmatrix}0 & B\\ 0 & 0\end{pmatrix},$$

並且恒有：$n\geqslant 2r(A)$。

7. 设n阶非异实方阵A满足：$A^{-1}=2I_n-A$，且I_n-A的秩为r，求证：

(i) 必存在非异阵P，使

$$P^{-1}AP=\begin{pmatrix}I_r & 0 & I_r\\ 0 & I_{n-2r} & 0\\ 0 & 0 & I_r\end{pmatrix};$$

(ii) A必可分解为两个(实)对称阵的乘积。

8. 用秩的降阶定理求下列矩阵的秩：

(i)
$$\begin{pmatrix}1 & -2 & 0 & 1 & 1 & -4\\ 2 & -4 & 3 & 2 & 0 & 2\\ 0 & 1 & -1 & 1 & 1 & -1\\ -1 & -1 & -3 & 0 & 4 & 2\end{pmatrix};$$

(ii) $(n-1)\times n$阵：

$$\begin{pmatrix}n+1 & 2 & 3\cdots n-1 & n\\ 1 & n+2 & 3\cdots n-1 & n\\ \cdots & \cdots & \cdots & \cdots\\ 1 & 2 & 3\quad 2n-1 & n\end{pmatrix}.$$

9. 设A是n阶非异阵，B是$m\times n$阵，α与β都是n维列向量，求证：

(i) $r(A-B'B)+r(I_m-BA^{-1}B')=n-m$；

(ii) $r(A+\alpha\beta')\geqslant n-1$，而$r(A+\alpha'\beta)=n\Longleftrightarrow\beta'A^{-1}\alpha=-1$。

10. 设A是$2k$阶非异反对称阵，α是$2k$维(实)列向量，求证：

$$r\begin{pmatrix}A & \alpha\\ \alpha' & 0\end{pmatrix}=2k.$$

11. 求下面矩阵的满秩分解：

$$\begin{pmatrix}6 & 0 & 3 & 1\\ 4 & -2 & 6 & 0\\ 2 & 2 & -3 & 1\end{pmatrix}.$$

12. (Frobenius)证明 Sylvester 公式的推广公式：

$$r(ABC)\geqslant r(AB)+r(BC)-r(B)$$

其中 A、B、C 分别是 $m\times n$ 阵、$n\times l$ 阵、$l\times s$ 阵。

（提示：作 B 的满秩分解，再应用 Sylvester 不等式）

13. (i) 试用 A 与 B 的满秩分解证明：$r(A+B)\leqslant r(A)+r(B)$；

(ii) 证明：$r(A-B)\geqslant r(A)-r(B)$；

(iii) 当 A 与 B 都是方阵，证明：$r(AB-I)\leqslant r(A-I)+r(B-I)$；

(iv) 由 (i) 及 Sylvester 不等式证明：若 $A^2=I_n$，则

$$r(I_n-A)+r(I_n+A)=n.$$

14. 设 $A=HL$ 是秩为 r 的 n 阶阵 A 的满秩分解，求证：

(i) 当 A 是实对称阵时，则 LH 是 r 阶非异阵；

(ii) A 是幂等阵的充要条件为 $LH=I_r$；

(iii) A 是实对称、幂等阵的充要条件为 $A=H(H'H)^{-1}H'$，或 $A=L'(LL')^{-1}L$.

15. 设 $A_1, A_2, \cdots, A_s$ 都是 n 阶阵，且 $A_1+A_2+\cdots+A_s=I_n$，又设 A_i 的秩为 n_i，$i=1, 2, \cdots, s$，求证：

(i) 如果 $A_iA_j=0$，$i\neq j$，$i, j=1, 2, \cdots, s$，则 $A_1, A_2, \cdots, A_s$ 都是幂等阵，且 $\sum_{i=1}^{s} n_i=n$.

(ii) 如果 $\sum_{i=1}^{s} n_i=n$，则 $A_iA_j=\delta_{ij}A_j$，$i, j=1, 2, \cdots, s$.

（提示：作 A_i 的满秩分解：$A_i=H_iL_i$，其中 H_i 与 L_i 分别是 $n\times n_i$ 列满秩阵与 $n_i\times n$ 行满秩阵，$i=1, 2, \cdots, s$。由假设，$\sum_{i=1}^{s} A_i=I_n$ 及 $A_iA_j=0$，$i\neq j$ 即可证得 $A_iA_j=\delta_{ij}A_j$，由此可证：$L_iH_j=\delta_{ij}I_{nj}$，然后推出 $\sum_{i=1}^{s} n_i=n$. 这证明了 (i)。

反之，由 $\sum_{i=1}^{s} A_i=I_n$，可证

$$(H_1, H_2, \cdots, H_s)\begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_s \end{pmatrix}=I_n,$$

再由假设 $\sum_{i=1}^{s} n_i=n$ 进一步可证 $(H_1, H_2, \cdots, H_s)$ 是 n 阶非异阵，由此可证得 $L_iH_j=\delta_{ij}I_{nj}$，故 $A_iA_j=\delta_{ij}A_j$，$i, j=1, 2, \cdots, s$.）

选 做 题

1. 求证：n 维列向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关的充要条件是，$|G'G|\neq$

0，其中 $G=(\alpha_1, \alpha_2, \cdots, \alpha_r)$。

*2. 设 $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$ 与 $\beta_1, \beta_2, \cdots, \beta_{n-r}$ 是 $m\times n$ 齐次方程组 $Ax=0$ 的两个基础解系，此处 $r(A)=r<n$，求证：必存在 $n-r$ 阶非异阵 P，使 $(\beta_1, \beta_2, \cdots, \beta_{n-r})=(\alpha_1, \alpha_2, \cdots, \alpha_{n-r})P$。

3. 求证：$m\times n$ 方程组 $Ax=b$ 相容的充要条件是，方程组 $A'y=0$ 的任何解 β 都满足 $\beta'b=0$。

（提示：必要性显然，充分性，由 $r(A)=r(A')=r$，取 $\beta_1, \beta_2, \cdots, \beta_{n-r}$ 为 $A'y=0$ 的基础解系，则 $\begin{pmatrix}A'\\b'\end{pmatrix}(\beta_1, \beta_2, \cdots, \beta_{n-r})=0$，于是由§4例3便知，

$r\begin{pmatrix}A'\\b'\end{pmatrix}\leqslant r$，所以 $r(A, b)=r\begin{pmatrix}A'\\b'\end{pmatrix}=r=r(A)$。）

4. 如果 $m\times n$ 方程组 $Ax=0$ 的解是方程 $b_1x_1+b_2x_2+\cdots+b_nx_n=0$ 的解，$x=(x_1, x_2, \cdots, x_n)'$，求证 $(b_1, b_2, \cdots, b_n)$ 必是 A 的 m 个行向量的线性组合。

5. 设 A 与 B 分别是 $m\times n$ 阵与 $n\times p$ 阵，

(i) 求证：方程组 $ABx=0$ 与方程组 $Bx=0$ 同解的充要条件是，$r(AB)=r(B)$；

(ii) $\overline{A}'Ax=0$ 与 $Ax=0$ 必是同解方程组，试证之；

(iii) 求证：$r(\overline{A}'A)=r(A)$，$r(A\overline{A}')=r(A)$。

6. 设 A 与 B 都是 n 阶对合阵，且 $|AB|<0$，求证：必存在 n 维非零列向量 ξ，使 $(-BA)\xi=\xi$。

7. (Ballantine, 1978) 设 A 是 n 阶奇异阵，且 $A=B_1B_2\cdots B_k$，其中 $B_1, B_2, \cdots, B_k$ 都是幂等阵，求证：$r(I_n-A)\leqslant k(n-r(A))$。

（提示：先有 $I_n-A=I_n-B_1B_2\cdots B_{k-1}B_k=(I_n-B_1)+(B_1-B_1B_2)+(B_1B^2-B_1B_2B_3)+\cdots+(B_1B_2\cdots B_{k-1}-B_1B_2\cdots B_{k-1}B_k)$，然后应用§4例2的性质，即(5)式，再应用定理1.6。）

8. 设 n 阶(实)对称阵(或反对称阵) A 的 r 阶主子式 $D\neq 0$，而 D 的所有 $r+1$ 阶、$r+2$ 阶加边主子式(即包含 D 的 $r+1$ 阶、$r+2$ 阶主子式)全等于零。求证：$r(A)=r$。

（提示：设 $\begin{pmatrix}D & \alpha\\ \beta & a_{ii}\end{pmatrix}$ 是包含 D 的 $r+1$ 阶阵，作 $r+2$ 阶主子阵：

$$M=\begin{pmatrix}D & \alpha & \beta'\\ \alpha' & a_{ii} & a_{is}\\ \beta & a_{si} & a_{ss}\end{pmatrix},$$

然后证明，$\left|(\mathrm{adj}M)\begin{pmatrix} s & t \\ s & t \end{pmatrix}\right|=0$，从而证明$\begin{vmatrix} D & \alpha \\ \beta & a_{st} \end{vmatrix}=0$。对反对称阵可仿此证明。）

9．设 n 阶（实）对称阵（反对称阵）的所有 $r+1$ 阶主子式、$r+2$ 阶主子式全等于零，求证：A 的任何 $r+1$ 阶子式等于零。

10．设 A 是 n 阶实方阵，且 $AA'=pI_n$，$p>0$，求证：

$$\mathrm{r}\left(A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right)=2k-n+\mathrm{r}\left(A_c\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right).$$

（提示：应用秩的升阶公式）

11．设 A 是 n 阶实方阵，且 $A^2=pI_n$，$p\neq 0$，求证：

$$\mathrm{r}\left(A\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right)=2k-n+\mathrm{r}\left(A_c\begin{pmatrix} j_1 & j_2\cdots j_k \\ i_1 & i_2\cdots i_k \end{pmatrix}\right).$$

12．设 A 是 n 阶非异阵，求证

$$\mathrm{r}\left((\mathrm{adj}A)\begin{pmatrix} i_1 & i_2\cdots i_k \\ j_1 & j_2\cdots j_k \end{pmatrix}\right)=2k-n+\mathrm{r}\left(A_c\begin{pmatrix} j_1 & j_2\cdots j_k \\ i_1 & i_2\cdots i_k \end{pmatrix}\right).$$

13．(i) 设$\left|A\begin{pmatrix} 1 & 2\cdots k \\ 1 & 2\cdots k \end{pmatrix}\right|\neq 0$ 是 $m\times n$阵 A 的 k 阶顺序主子式，记 $s_{ij}=\left|A\begin{pmatrix} 1 & 2\cdots k & i \\ 1 & 2\cdots k & j \end{pmatrix}\right|$，$i=k+1,\cdots,m$；$j=k+1\cdots,n$，且

$$S=\begin{pmatrix} S_{k+1,k+1}\cdots S_{k+1,n} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ S_{m,k+1}\quad\cdots S_{mn} \end{pmatrix},$$

称 S 为 **Sylvester 矩阵**，求证：

$$\mathrm{r}(A)=k+\mathrm{r}(S),$$

称上式为矩阵的**秩的第三降阶定理**。

(ii) 若 $a_{11}\neq 0$，求证：

$$\mathrm{r}\begin{pmatrix} a_{11} & a_{12} & a_{13}\cdots a_{1n} \\ a_{21} & a_{22} & a_{23}\cdots a_{2n} \\ \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3}\cdots a_{mn} \end{pmatrix}=1+\mathrm{r}\begin{pmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1n} \\ a_{21} & a_{2n} \end{vmatrix} \\ \cdots & \cdots & \cdots & \cdots \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{m1} & a_{m2} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{m1} & a_{m3} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1n} \\ a_{m1} & a_{mn} \end{vmatrix} \end{pmatrix}.$$

（提示：应用秩的第二降阶定理与行列式的第一降阶定理，並仿照行列式的第三降阶定理的证明即得(i)．由(i)显然可得(ii)。）

14. 求证：$m\times n$ 阵 A 的秩为 r 的充要条件是，A 有分解式：

$$A=\alpha_1\beta_1'+\alpha_2\beta_2'+\cdots+\alpha_r\beta_r'$$

其中，$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r$ 为线性无关的 m 维列向量；$\beta_1,\ \beta_2,\ \cdots,\ \beta_r$ 是线性无关的 n 维列向量。

（提示：应用 A 的满秩分解）

15. 设 $A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{m\times n}$，则称 $m\times n$ 阵：

$$A*B=(a_{ij}b_{ij})_{m\times n}=\begin{pmatrix}a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n}\\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn}\end{pmatrix}$$

为 A 与 B 的 **Hadamard 乘积**。

(i) 若 $A,\ B_1,\ B_2,\ \cdots,\ B_s$ 都是 $m\times n$ 阵，求证：

$$A*\left(\sum_{i=1}^{s}B_i\right)=\sum_{i=1}^{s}(A*B_i);$$

(ii) 若 A、B 都是以数为元素的 $m\times n$ 阵，求证：$A*B=B*A$；

(iii) 若 α 与 β 都是 m 维列向量；γ 与 δ 都是 n 维列向量，求证：$(\alpha*\beta)(\gamma*\delta)'=(\alpha\gamma')*(\beta\delta')$；

(iv) (Ballantine, 1968)　设 A 与 B 都是 $m\times n$ 阵，求证：

$$\mathrm{r}(A*B)\leqslant\mathrm{r}(A)\cdot\mathrm{r}(B).$$

（提示：应用本题(i)、(iii)以及第 14 题）

16. (i) 设 A 是秩为 r 的 n 阶阵，求证：必存在 n 阶非异阵 P，使 $P'AP=\begin{pmatrix}B\\0\end{pmatrix}$，其中 B 为 $r\times n$ 行满秩阵，

(ii) 对秩为 r 的 n 阶（实）对称阵，求证：必存在 n 阶非异阵 P，使 $P'AP=\begin{pmatrix}S & 0\\0 & 0\end{pmatrix}$，其中 S 为 r 阶非异对称阵；

(iii) 求证：秩为 r 的 n 阶(实)对称阵 A 必有分解式，$A=HSH'$，其中 H 是 $n\times r$ 列满秩阵，S 是 r 阶非异阵（因而 SH' 是 $r\times n$ 行满秩阵）。

17. 设 A 是秩为 r 的 n 阶对称阵，求证：对任何 $1\leqslant i_1<i_2<\cdots<i_r\leqslant n$，$1\leqslant j_1<j_2<\cdots<j_r\leqslant n$，必有

$$\left|A\begin{pmatrix}i_1 & i_2\cdots i_r\\ j_1 & j_2\cdots j_r\end{pmatrix}\right|^2=\left|A\begin{pmatrix}i_1 & i_2\cdots i_r\\ i_1 & i_2\cdots i_r\end{pmatrix}\right|\cdot\left|A\begin{pmatrix}j_1 & j_2\cdots j_r\\ j_1 & j_2\cdots j_r\end{pmatrix}\right|.$$

（提示：先写

$$A\begin{pmatrix} i_1 & i_2\cdots i_r \\ j_1 & j_2\cdots j_r \end{pmatrix}=\begin{pmatrix} e'_{i_1} \\ e'_{i_2} \\ \vdots \\ e'_{i_r} \end{pmatrix}A(e_{j_1}, e_{j_2}, \cdots, e_{j_r}),$$

然后应用 16 题的(iii))。

18．由 17 题证明：

(i) 秩为 r 的 n 阶(实)对称阵 A 至少有一个 r 阶主子式不等于零；

(ii) 如果秩为 r 的 n 阶(实)对称阵 A 至少有一个 r 阶主子式等于零，则 A 至少有 C_n^r-1 个 r 阶子式等于零；

(iii) 如果秩为 r 的 n 阶(实)对称阵 A 的所有 r 阶主子式全不等于零，则 A 的所有 r 阶子式都不等于零。

19．设秩为 r 的 n 阶(实)对称阵 A 的 r 阶主子阵 $A_r=A\begin{pmatrix} i_1 & i_2\cdots i_r \\ i_1 & i_2\cdots i_r \end{pmatrix}$ 是非异阵，

(i) 设 $T=P_{i_1\cdots i_r i_{r+1}\cdots i_n}=(e_{i_1}, e_{i_2}, \cdots, e_{i_r}, e_{i_{r+1}}, \cdots, a_{in})$ 是 n 阶排列阵，$1\leqslant i_{r+1}<\cdots<i_n\leqslant n$，记

$$B=\begin{pmatrix} e'_{i_{r+1}} \\ e'_{i_{r+2}} \\ \vdots \\ e'_{i_n} \end{pmatrix}A(e_{i_1}, e_{i_2}, \cdots, e_{i_r}),\ H=P\begin{pmatrix} I_r \\ BA_r^{-1} \end{pmatrix},$$

求证：A 必有满秩分解：$A=HA\begin{pmatrix} i_1 & i_2\cdots i_r \\ i_1 & i_2\cdots i_r \end{pmatrix}H'$；

(ii) 求证：对任一 r 阶子矩阵 $A\begin{pmatrix} j_1 & j_2\cdots j_r \\ k_1 & k_2\cdots k_r \end{pmatrix}$，必存在 $n\times r$ 列满秩阵 H，使

$$A\begin{pmatrix} j_1 & j_2\cdots j_r \\ k_1 & k_2\cdots k_r \end{pmatrix}=H\begin{pmatrix} j_1 & j_2\cdots j_r \\ 1 & 2\cdots r \end{pmatrix}A\begin{pmatrix} i_1 & i_2\cdots i_r \\ i_1 & i_2\cdots i_r \end{pmatrix}\left(H'\begin{pmatrix} 1 & 2\cdots r \\ k_1 & k_2\cdots k_r \end{pmatrix}\right).$$

20．求证：第 16 题到第 19 题的所有结论对秩为 $r=2p$ 的 n 阶反对称阵都成立。

21．设 A 是秩为 r 的 $m\times n$ (实)阵，如果存在 $n\times m$ (实)阵 B，使

(i) AB 是实对称阵；

(ii) BA 是实对称阵；

(iii) $ABA=A$；

(iv) $BAB=B$，

则称 B 是 A 的广义逆(矩阵)，求证：

(i) A 若有广义逆矩阵 B，则 B 是 (由 A) 唯一(确定)的，故记为 $B=A^+$；

(ii) 对 $m\times r$ 列满秩阵 H 以及 $r\times n$ 行满秩阵 L，它们的广义逆均存在，且

$$H^+=(H'H)^{-1}H',\ L^+=L'(LL')^{-1};$$

(iii) 如果 $A=HL$ 是秩为 r 的 $m\times n$ 阵 A 的满秩分解，则

$$A^+=L'(LL')^{-1}(H'H)^{-1}H';$$

(iv) 任何非异阵都有广义逆。

(提示：(i) 可反复应用广义逆定义的四条，设 B 与 C 都是 A 的广义逆，则可证得 $B=BAC$，$C=BAC$，故 $B=C$。)

22. 设 A^+ 是 $m\times n$ 阵 A 的广义逆，求证：

(i) $r(A^+)=r(AA^+)=r(A^+A)=r(A)$；

(ii) $r(I_n-A^+A)=n-r(A)$；

(iii) $m\times n$ 方程组 $Ax=b$ 相容的充要条件是，$AA^+b=b$。

(提示：对(i)，易证：$A^+=L^+H^+$，$H^+H=I_r=LL^+$，再用 Sylvester 不等式的推论即可，对(ii)，应用矩阵的秩的第二降阶定理及 $A^+=L^+H^+$ 即可。)

第四章　多　项　式

本章主要介绍与一元高次方程密切相关的一元多项式理论，重点是数域上一元多项式的分解因式定理与应用以及它们在复数域上的相应结果。

§1　集合、数环与数域

一、集合

集合是数学中常用的一个基本概念。集合，简称集，就是指一些事物的全体。例如，整数全体构成整数集；0 与 1 这两个数也构成一个数集；n 阶方阵全体构成一个方阵集合；线性方程组 $AX=b$ 的所有解构成它的解向量集合等等。集合通常以大写字母 S、T、M、N……记之。构成集合的事物称为该集合的**元素**或**元**，通常以小写字母 a、b、c……记之。

所谓给出一个集合，就是要指明它由哪些元素构成。一般有以下两种表示法：其一，列出集合中所有的元素(如果可能的话)，此法称为**枚举法**。例如，整数集 Z 可表示为

$$Z=\{0,\ \pm 1,\ \pm 2,\ \cdots,\ \pm n,\ \cdots\};$$

n 维向量 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 所成的集合可表示为 $\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$。其二，指出集合中元素所具有的特性。记号 $S=\{x\,|\,x$ 所具有的性质 $P\}$，表明集合 S 由所有具有性质 P 的元素组成，此法称为**综合法**。例如，$T=\{x\,|\,2x^2+3x+1=0,\ x$ 为复数$\}$代表方程 $2x^2+3x+1=0$ 的复根集合；$S=\{A\,|\,|A|=1,\ A$ 为 n 阶实方阵$\}$ 表示所有行列式等于 1 的 n 阶实方阵构成的集合；$U=\{(x,\ y)\,|\,x^2+y^2=1\}$表示圆心在原点半径为 1 的圆上点的全体所成的集合。

对于某个集合 S 来说，某一事物 a 或者是集合 S 中的元素，或若不是，两者必居其一。若 a 是 S 的元素，则记为 $a\in S$，读作 a 属于 S；否则，记为 $a\overline{\in}S$，读作 a 不属于 S。例如，

$$1\overline{\in}\{x\,|\,2x^2+3x+1=0,\ x\text{ 为复数}\};$$

n 阶单位阵 $I_n\in\{A\,|\,|A|=1,\ A\text{ 为 }n\text{ 阶实方阵}\}$。

为表达方便起见，称不含有任何元素的集合为**空集**。以 $\varnothing$ 记之。例如 $\{x\,|\,x>1\text{ 且 }x<1\}=\varnothing$；$\{x\,|\,2x^2+3x+3=0,\ x\text{ 为实数}\}=\varnothing$。若 T 不是空集，则记为 $T\neq\varnothing$，读作 T 非空。例如 $T=\{0\}$ 不是空集；$\{A\,|\,|A|=1,\ A\text{ 为实方阵}\}\neq\varnothing$。

设 S、T 是两个集合，称 S 为 T 的**子集**（记为 $S\subseteq T$），如果 S 中任一元素都是 T 的元素；称 S 为 T 的**真子集**（记为 $S\subset T$），如果 S 是 T 的子集，且 T 中至少有一个元素不属于 S。记号 $S\subseteq T(S\subset T)$ 读作 S 包含（真包含）在 T 中。规定空集是任何集合的子集。今后，恒以 Z、Q 表示整数集和有理数集，于是有 $\varnothing\subset Z\subset Q$。

称集合 S 与 T 相等（记为 $S=T$），如果 $S\subseteq T$ 且 $T\subseteq S$。显然，相等的两个集合含有相同的元素。称集合为**无限集**，如果其中含有无限多个不同的元素；否则，称为**有限集**。

集合之间进一步的关系，将在第九章中介绍。

二、数环与数域

数集是常见的一种集合。自然数集 N、整数集 Z、有理数集 Q、实数集 R、复数集 C 已为大家所熟知。

在数集 S 中，任取两数作某种运算，若其结果仍在 S 中，则称数集 S 对于这种运算是封闭的。显然，Z、Q、R、C 对于数的减法是封闭的，N 却不然；Q、R、C 对于数的除法（除数非零）是封闭的，Z 却不然。

代数中讨论问题时，经常涉及某一数集中数的加、减、乘、除运算及其性质。为了统一处理具有共同性质的数集，引进数环与数域的概念。

定义 称非空数集 S 为**数环**，如果 S 对于数的加、减、乘运算是封闭的，即

$$\forall a、b\in S\Longrightarrow a\pm b、ab\in S)$$①.

称至少包含两个不同元素的数集K为**数域**，如果K对于数的加、减、乘、除(除数非零)运算是封闭的，即

$$\forall a、b\in S\Longrightarrow a\pm b、ab、\frac{a}{b}(b\neq 0)\in S.$$

由定义可知，若某一数集是数域，则必是数环。Q、R和C是数域，分别称为**有理数域**、**实数域**和**复数域**。Z是数环，但不是数域，称为**整数环**。N不是数环，当然也不是数域。

例1 验证数集 $Q(\sqrt{2})=\{a+b\sqrt{2},\ \forall a、b\in Q\}$ 为一数域。

【证明】 首先，$0=0+0\sqrt{2}$，$1=1+0\sqrt{2}$，所以 $0、1\in Q(\sqrt{2})$。其次，$Q(\sqrt{2})$对数的加、减运算显然是封闭的。今取$Q(\sqrt{2})$中任意两数$a+b\sqrt{2}$与$c+d\sqrt{2}$，由于

$$(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2},$$

而当 $c+d\sqrt{2}\neq 0$ 时，必有 $c-d\sqrt{2}\neq 0$，且

$$\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}=\frac{ac-2bd}{c^2-2d^2}+\frac{bc-ad}{c^2-2d^2}\sqrt{2}.$$

又因为$a、b、c、d\in Q$，Q为数域，所以$ac+2bd\in Q$，$ad+bc\in Q$，即 $Q(\sqrt{2})$对于乘法是封闭的；注意到 $c^2-2d^2\neq 0$，所以

$$\frac{ac-2bd}{c^2-2d^2}\in Q,\quad \frac{bc-ad}{c^2-2d^2}\in Q,$$

即$Q(\sqrt{2})$对于除法也封闭。故$Q(\sqrt{2})$是一数域。 证毕。

类似地，可以验证数集

$$Z[\sqrt{2}]=\{x\mid x=a+b\sqrt{2},\ \forall a、b\in Z\}$$

为一数环。必须注意 $Z[\sqrt{2}]$不是数域。事实上，若在$Z[\sqrt{2}]$中特别取$x=\sqrt{2}$，$y=2$，则$\frac{x}{y}=\frac{1}{2}\sqrt{2}\overline{\in} Z[\sqrt{2}]$，即$Z[\sqrt{2}]$对于除法运算不封闭。

① 记号"∀…"的含义是"对所有的……"。

由定义不难得到下列性质。

命题 1 任一数环必包含零；任一数域必包含零与 1。

命题 2 有理数域是最小的数域。

命题 1 的证明留给读者，利用它的逆否命题立即可以断定：奇数集不是数环；偶数集不是数域。

下面证明命题 2。设 K 为任一数域。由命题 1，$1\in K$。利用 K 对于加法的封闭性，可得任一正整数 $n\in K$。注意到 $0\in K$ 并利用 K 对于减法的封闭性，可得任一负整数 $(-n)\in K$。综合之，$Z\subseteq K$。今任取 $x\in Q$，可设 $x=\frac{p}{q}$，其中 p、$q\in Z$，$q\neq 0$。因 $Z\subseteq K$，故 p、$q\in K$。因数域 K 对于除法运算封闭，所以 $x=\frac{p}{q}\in K$。这就证明了 $Z\subseteq K$。 证毕。

设 K_1、K_2 为两个数域，若 $K_1\subset K_2$，则称 K_1 为 K_2 的**真子域**，而称 K_2 为 K_1 的**扩域**。例如，R 是 C 的真子域，而 C 是 R 的扩域(因 $\sqrt{-1}\notin R$)。

习　题

1. 下列数集中，哪些是数域，哪些仅是数环，哪些不是数环？(说明理由)

(i) 非整数的有理数集合 S；

(ii) 偶数与分子为偶数的既约分数集合 T；

(iii) $Q(\sqrt{-1})=\{a+b\sqrt{-1},\forall a、b\in Q\}$；

(iv) $Q(\sqrt[3]{2})=\{a+b\sqrt[3]{2},\forall a、b\in Q\}$。

2. 设 M 为至少包含一个非零元的数集。

(i) 求证，若 M 对于数的减法、除法(除数非零)封闭，则对于数的加法、乘法也封闭；

(ii) 举例说明(i)中结论之逆未必成立。

3. (i) 数域 $Q(\sqrt{2})$ 是否为实数域 R 的真子域？

(ii) 数域 $Q(\sqrt{-1})$ 是否为实数域 R 的扩域？

§2 一元多项式

一、定义，有关的约定及名词

定义 设K为一数域，x为一符号，且$\forall a\in K$，形式积$ax^i=x^ia$（i为任一正整数）。若$a_0, a_1, \cdots, a_n\in K$，$n$为一非负整数，则称形式表达式

$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \tag{1}$$

为数域K上(关于x)的**一个一元多项式**。称x为**未定元**，称K为**基域**。

一元多项式通常以$f(x)$、$g(x)$、$h(x)\cdots$记之，或简记为f、g、$h\cdots$。约定非零常数a_0代表a_0x^0。称(1)中的a_ix^i为i**次项**，a_i为该项的**系数**；且当$a_i\neq0$时，称i为该项的**次数**($i=0, 1, \cdots, n$)。约定$0x^i$可以从多项式中删去；若多项式中不出现i次项，则可添上$0x^i$这一项。对于x^i，则认为其系数为1。我们将(1)式与表达式

$$a_0+a_1x+\cdots+a_{n-1}x^{n-1}+a_nx^n$$

不加区别。习惯上，今后多数采用"降幂(次)排列"式(1)，并常将它记为

$$\sum_{i=0}^{n}a_ix^i.$$

称多项式中系数非零的最高次项为它的**首项**。首项的次数称为多项式的次数。即若

$$f(x)=\sum_{i=0}^{n}a_ix^i,\ a_n\neq0,$$

则$f(x)$的次数为n，记为$\deg(f(x))=n$。特别地，非零常数是零次多项式。称各次项系数均为零的多项式为零多项式，以0记之。对于零多项式，无次数可言。因此，今后当且仅当$f(x)\neq0$时才书写记号$\deg(f(x))$。显然，$\deg(f(x))\geqslant0$。

二、多项式的运算及其性质

数域K上一元多项式全体构成的集合记为$K[x]$即

$$K[x]=\left\{\sum_{i=0}^{n}a_ix^i,\ \forall a_0, a_1, \cdots, a_n\in K, n\text{为任一非负整数}\right\}.$$

定义 设$f(x)$、$g(x)\in K[x]$，称$f(x)$与$g(x)$相等（记为

$f(x)=g(x)$)，如果它们同次项的系数全都对应相等。

由定义可知，次数不相同的多项式一定不相等；与零多项式相等的多项式只能是零多项式，即若

$$f(x)=\sum_{i=0}^{n}a_ix^i,$$

则

$$f(x)=0 \Longleftrightarrow a_i=0(i=0,\ 1,\ \cdots,\ n).$$

定义 $\forall f(x)$、$g(x)\in K[x]$，在其中适当添上一些系数为零的项，总可设

$$f(x)=\sum_{i=0}^{n}a_ix^i, \tag{2}$$

$$g(x)=\sum_{i=0}^{n}b_ix^i. \tag{3}$$

令 $h(x)=\sum_{i=0}^{n}(a_i+b_i)x^i$，显然，$h(x)\in K[x]$。称 $h(x)$ 为 $f(x)$ 与 $g(x)$相加所得的和，记为

$$f(x)+g(x)=\sum_{i=0}^{n}(a_i+b_i)x^i. \tag{4}$$

令 $-g(x)=\sum_{i=0}^{n}(-b_i)x^i$，显然，$-f(x)\in K[x]$。称 $-g(x)$ 为 $g(x)$ 的负元素，且称 $f(x)+(-g(x))$ 为 $f(x)$ 减去 $g(x)$ 所得的差，记为

$$f(x)-g(x)=f(x)+(-g(x)).$$

这里，减法是由加法来定义的。因此，对减法的讨论将归并到对加法的讨论之中去，而不另外列出。

不难验证，对于多项式的加法，有下列运算规则：

(i) $f(x)+g(x)=g(x)+f(x)$；

(ii) $(f(x)+g(x))+h(x)=f(x)+(g(x)+h(x))$；

(iii) $f(x)+0=f(x)$；

(iv) $f(x)+(-f(x))=0$。

下面仅证明(i)，其余留给读者作为练习。事实上，若 $f(x)$、$g(x)$ 如(2)、(3)所设，根据加法定义，$f(x)+g(x)$ 的第 i 次项的系数为

a_i+b_i; $g(x)+f(x)$的第 i 次项的系数为 b_i+a_i。注意到数的加法可交换。即有 $a_i+b_i=b_i+a_i(i=0, 1, \cdots, n)$，再由多项式相等的定义，可得 $f(x)+g(x)=g(x)+f(x)$。

由加法运算规则(ii)，定义

$$f_1(x)+f_2(x)+\cdots+f_{n-1}(x)+f_n(x)$$
$$=(f_1(x)+f_2(x)+\cdots+f_{n-1}(x))+f_n(x),$$

其中 $n\geqslant 3$。应用归纳法可以证明：对任何 k，$1\leqslant k\leqslant n-1$，成立

$$f_1(x)+f_2(x)+\cdots+f_n(x)$$
$$=(f_1(x)+f_2(x)+\cdots+f_k(x))+(f_{k+1}(x)+\cdots+f_n(x)),$$

称上式为加法的**一般结合律**(留作练习)。

在多项式的运算中，加法消去律是常用的。

命题 1 设 $f(x)$、$g(x)$、$h(x)\in K[x]$，且

$$f(x)+h(x)=g(x)+h(x),$$

则

$$f(x)=g(x).$$

【证明】 因 $f(x)+h(x)=g(x)+h(x)$，故有

$$f(x)+h(x)+(-h(x))=g(x)+h(x)+(-h(x)),$$

利用加法运算规则(ii)、(iv)、(iii)即可得到 $f(x)=g(x)$。

证毕。

推论 设 $f(x)$、$g(x)\in K[x]$，则

$$f(x)-g(x)=0 \Longleftrightarrow f(x)=g(x).$$

【证明】 注意到命题 1 的逆命题显然成立，于是有

$$f(x)+h(x)=g(x)+h(x) \Longleftrightarrow f(x)=g(x).$$

上式对任意的 $f(x)$、$g(x)$、$h(x)\in K[x]$ 都成立，特别地，令

$$h(x)=-g(x),$$

并利用减法的定义及加法运算规则(iv)即可得证。 证毕。

定义 $\forall f(x)$、$g(x)\in K[x]$，设

$$f(x)=\sum_{i=0}^{n}a_ix^i,\quad g(x)=\sum_{j=0}^{m}b_jx^j,$$

令 $h(x)=\sum\limits_{r=0}^{n+m}(a_rb_0+a_{r-1}b_1+\cdots+a_0b_r)x^r$，显然，$h(x)\in K[x]$，

称 $h(x)$ 为 $f(x)$ 与 $g(x)$ 相乘所得的积，记为

$$f(x)g(x)=\sum_{r=0}^{n+m}\Big(\sum_{i+j=r}a_ib_j\Big)x^r. \tag{5}$$

例如，若 $f(x)=x^2+ax+b$，$g(x)=x^2+cx+d$，则

$$f(x)g(x)=x^4+(c+a)x^3+(d+ac+b)x^2+(ad+bc)x+bd.$$

又如，$\forall f(x)\in K[x]$，必定有 $0\cdot f(x)=0$。

不难验证，对于多项式的乘法，有下列运算规则：

(i) $f(x)g(x)=g(x)f(x)$；

(ii) $(f(x)g(x))h(x)=f(x)(g(x)h(x))$；

(iii) $(f(x)+g(x))h(x)=f(x)h(x)+g(x)h(x)$。

以上等式的证明与加法交换律(i)的证明类似，其主要一步是考察有关多项式的同次项系数是否全都对应相等，从而使问题转化为对于数来证明相应的运算规则。以(ii)式为例证之，若以 a_i、b_i、c_i 分别代表 $f(x)$、$g(x)$、$h(x)$的 i 次项系数，由乘法定义，$f(x)g(x)$中 r 次项的系数应为

$$\sum_{i+j=r}a_ib_j,$$

$(f(x)g(x))h(x)$ 中 S 次项的系数应为

$$\sum_{r+k=s}\Big(\sum_{i+j=r}a_ib_j\Big)c_k;$$

同理，$g(x)h(x)$ 中 t 次项的系数应为

$$\sum_{j+k=t}b_jc_k,$$

$f(x)(g(x)h(x))$中 S 次项的系数应为

$$\sum_{i+t=s}a_i\Big(\sum_{j+k=t}b_jc_k\Big).$$

利用数的运算规则(用了哪几条?)可得

$$\sum_{r+k=s}\Big(\sum_{i+j=r}a_ib_j\Big)c_k=\sum_{i+j+k=s}a_ib_jc_k,$$

$$\sum_{i+t=s}a_i\Big(\sum_{j+k=t}b_jc_k\Big)=\sum_{i+j+k=s}a_ib_jc_k,$$

即(ii)式两端多项式的同次项系数全都对应相等，故(ii)式成立。

由乘法运算规则(ii)，定义

$$f_1(x)f_2(x)\cdots f_{n-1}(x)f_n(x)=(f_1(x)f_2(x)\cdots f_{n-1}(x))f_n(x),$$

其中 $n\geqslant 3$。应用归纳法也可证明乘法的一般结合律：

$$f_1(x)f_2(x)\cdots f_n(x)$$
$$=(f_1(x)f_2(x)\cdots f_k(x))(f_{k+1}(x)\cdots f_n(x)),\ 1\leqslant k\leqslant n-1。$$

定义

$$f^n(x)=\begin{cases}1, & \text{当 } n=0 \text{ 时;}\\ f^{n-1}(x)f(x), & \text{当 } n\geqslant 1 \text{ 时。}\end{cases}$$

称 $f^n(x)$ 为多项式 $f(x)$ 的幂，其中 n 为非负整数。

与数的乘法相似，多项式乘法亦无零因子，因而消去律成立。

命题 2 若 $f(x)$、$g(x)\in K[x]$，且 $f(x)$、$g(x)$ 全不为零，则

$$f(x)g(x)\neq 0。$$

【证明】 因 $f(x)$、$g(x)$ 全不为零，不妨设它们的首项分别为 a_nx^n、b_mx^m。由乘法定义，$f(x)g(x)$ 的最高次项为 $a_nb_mx^{n+m}$。注意到 $a_n\neq 0$，$b_m\neq 0$，故 $a_nb_m\neq 0$，即可得 $f(x)g(x)\neq 0$。　　证毕。

推论 若 $f(x)h(x)=g(x)h(x)$，且 $h(x)\neq 0$，则

$$f(x)=g(x)。$$

【证明】 因 $f(x)h(x)=g(x)h(x)$，利用命题 1 的推论及乘法运算规则(iii)可得

$$(f(x)-g(x))h(x)=0。$$

又因 $h(x)\neq 0$，由命题 2 即得 $f(x)-g(x)=0$，从而 $f(x)=g(x)$。

证毕。

前已约定，任一非零常数代表零次多项式。在 (5) 式中令 $f(x)=c$，$c\neq 0$，可得

$$cg(x)=\sum_{j=0}^{m}(cb_j)x^j,$$

称 $cg(x)$ 为 c 与 $g(x)$ 相乘(数乘)所得的积。

多项式的数乘，作为其乘法的特款，有下列运算规则：

(i) $(kl)f(x)=k(lf(x))$；

(ii) $(k+l)f(x)=kf(x)+lf(x)$；

(iii) $k(f(x)+g(x))=kf(x)+kg(x)$；

(iv) $1\cdot f(x)=f(x)$。

命题 3 设 $f(x)$、$g(x)\in K[x]$，则

(i) $\deg(f(x)\pm g(x))\leqslant\max(\deg(f(x)),\deg(g(x)))$①；

(ii) $\deg(f(x)g(x))=\deg(f(x))+\deg(g(x))$。

事实上，由加减法定义立即可得 (i)。(ii) 的证明已包含在命题 2 的证明之中。

最后，我们指出，$K[x]$ 中按 (4)、(5) 定义了加法、乘法之后，称为**一元多项式环**. 读者稍加注意就会发现，若将 K 改为包含 1 的数环，本节所有的讨论结果仍然成立。

习 题

1. 证明多项式的下列运算规则：

(i) $(f(x)+g(x))+h(x)=f(x)+(g(x)+h(x))$；

(ii) $(f(x)+g(x))h(x)=f(x)h(x)+g(x)h(x)$；

(iii) $(kl)f(x)=k(lf(x))$；

(iv) $f(x)(lg(x))=lf(x)g(x)$；

(v) $(-1)f(x)=-f(x)$。

2. 由等式 $(1+x)^m(1+x)^n=(1+x)^{m+n}$ 证明

$$c_m^k+c_m^{k-1}c_n^1+\cdots+c_m^1c_n^{k-1}+c_n^k=c_{m+n}^k.$$

3. 设多项式 $f(x)=q(x)g(x)+r(x)$，且

$$\deg(r(x))<\deg(g(x))\leqslant\deg(f(x)),$$

证明，

$$\deg(q(x))=\deg(f(x))-\deg(g(x)).$$

4. 证明关于多项式加法的一般结合律。

§3 整 除 性

以 K 代表某一数域，下面的讨论在多项式环 $K[x]$ 中进行。

① 记号 $\max(a,b)$ 表示 a,b 之中较大的一个数。

一、整除性及其性质

两个多项式相除未必都能除尽，于是有下面关于整除的定义。

定义 设 $f(x)$、$g(x)\in K[x]$，称 $g(x)$能整除 $f(x)$（记为 $g|f$），如果存在 $q(x)\in K[x]$，使得

$$f(x)=q(x)g(x);$$

否则，称 $g(x)$ 不能整除 $f(x)$（记为 $g\nmid f$）。

当 $g|f$ 时，也称 $f(x)$ 能被 $g(x)$ 整除，此时称 $g(x)$ 为 $f(x)$ 的**因式**，$f(x)$ 为 $g(x)$ 的**倍式**。记号

$$\frac{f(x)}{g(x)}$$

表示 $g|f$，且 $g(x)\neq 0$。

由定义可得下列性质：

(i) 零多项式的因式可以是任一多项式，但其倍式只能是零多项式；

(ii) 任一多项式 $f(x)$ 必以 c、$cf(x)$ 为因式，其中 c 为任一个非零常数；

(iii) 若 $f(x)|g(x)$，$g(x)|h(x)$，则 $f(x)|h(x)$；

(iv) 若 $f(x)|g_i(x)(i=1,2,\cdots\cdots,s)$，则

$$\forall u_i(x)\in K[x](i=1,2,\cdots,s),$$

成立 $f(x)\left|\sum_{i=1}^{s}u_i(x)g_i(x)\right.$；

(v) $f(x)|g(x)$ 且 $g(x)|f(x)\Longleftrightarrow f(x)=cg(x)$，$c\neq 0$。

下面仅证明性质(v)的必要性，其余的证明是容易的。由条件可设

$$f(x)=h_1(x)g(x),\ g(x)=h_2(x)f(x),$$

于是

$$f(x)=h_1(x)h_2(x)f(x)。$$

若 $f(x)=0$，则 $g(x)=h_2(x)f(x)=0$，从而 $f(x)=g(x)$，结论成立；若 $f(x)\neq 0$，由消去律可得

$$h_1(x)h_2(x)=1,$$

上式蕴含 $h_1(x)\neq 0$，$h_2(x)\neq 0$。据§2 中命题 3，应有

$$\deg(h_1(x))+\deg(h_2(x))=0,$$

从而 $\deg(h_1(x))=\deg(h_2(x))=0$。故可得 $h_1(x)=c$，$c\neq0$，(v) 的必要性得证。

二、带余除法

定理 3.1 设 $f(x)$ 是 $K[x]$ 中任一多项式，则对 $K[x]$ 中的多项式 $g(x)\neq0$，必存在 $q(x)$、$r(x)\in K[x]$，使得

$$f(x)=q(x)g(x)+r(x),$$

其中，$r(x)=0$ 或者 $\deg(r(x))<\deg(g(x))$；且这样的 $q(x)$、$r(x)$ 由 $f(x)$、$g(x)$ 唯一确定，分别称它们为 $g(x)$ 除 $f(x)$ 的**商式**与**余式**。

【证明】 将中学代数课程中的长除法一般化，即可获得存在性的证明。下面用第二数学归纳法来表达。

若 $f(x)=0$ 或 $\deg(f)<\deg(g)$，则取 $q(x)=0$，$r(x)=f(x)$ 即可；

若 $\deg(f)\geqslant\deg(g)$，则对 $\deg(f)$ 作归纳。设

$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0,\ b_m\neq0.$$

当 $\deg(f)=0$ 时，设 $f(x)=a_0$，$a_0\neq0$，此时 $g(x)=b_0$，$b_0\neq0$。注意到

$$f(x)=(a_0b_0^{-1})b_0,$$

取 $q(x)=a_0b_0^{-1}$，$r(x)=0$ 即可。当 $\deg(f)=n$ 时，设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\ a_n\neq0.$$

因 $n\geqslant m$ 且 $b_m\neq0$，所以 $a_nb_m^{-1}x^{n-m}\in K[x]$。令

$$f_1(x)=f(x)-a_nb_m^{-1}x^{n-m}g(x),$$

显然，$f_1(x)\in K[x]$ 且 $f_1(x)=0$ 或者 $\deg(f_1)<\deg(f)=n$。若

$$f_1(x)=0\ \text{或}\ \deg(f_1)<\deg(g),$$

则取 $q(x)=a_nb_m^{-1}x^{n-m}$，$r(x)=f_1(x)$ 即可；若 $\deg(f_1)\geqslant\deg(g)$，注意到 $\deg(f_1)<n$，由归纳假设，必存在 $q_1(x)$、$r_1(x)\in K[x]$，使得

$$f_1(x)=q_1(x)g(x)+r_1(x),$$

其中，$r_1(x)=0$ 或者 $\deg(r_1)<\deg(g)$。取

$$q(x)=a_n b_m x^{n-m}+q_1(x),\ r(x)=r_1(x)$$
即可。

再证唯一性。设另有 $\tilde{q}(x)$、$\tilde{r}(x)\in K[x]$，使得

$$f(x)=\tilde{q}(x)g(x)+\tilde{r}(x),$$

其中，$\tilde{r}(x)=0$ 或者 $\deg(\tilde{r})<\deg(g)$。由

$$q(x)g(x)+r(x)=\tilde{q}(x)g(x)+\tilde{r}(x),$$

可得

$$\tilde{r}(x)-r(x)=(q(x)-\tilde{q}(x))g(x).$$

若 $q(x)\neq\tilde{q}(x)$，则 $\tilde{r}(x)\neq r(x)$。又 $g(x)\neq 0$，由 § 2 的命题 3 得到

$$\deg(\tilde{r}-r)=\deg(q-\tilde{q})+\deg(g)\geqslant\deg(g).$$

但此时 $r(x)$、$\tilde{r}(x)$不全为零，故只可能有

$\deg(\tilde{r}-r)=\deg(\tilde{r})<\deg(g)$（当 $r(x)=0$，$\tilde{r}(x)\neq 0$ 时）或者

$$\deg(\tilde{r}-r)\leqslant\max(\deg(\tilde{r}),\deg(r))<\deg(g).$$

这就产生了矛盾，从而必有

$$q(x)=\tilde{q}(x),\ r(x)=\tilde{r}(x).$$

证毕。

定理 3.1 是一元多项式环中的一个基本性质，通常称之为一元多项式环中有带余除法，它是下面讨论最大公因式及分解因式的基础。

定理 3.1 的证明给出了求 $q(x)$、$r(x)$ 的具体方法。进一步，便可获得一个判别整除的方法。

定理 3.2 设 $f(x)$、$g(x)\in K[x]$，$g(x)\neq 0$，则 $g(x)$ 能整除 $f(x)$ 的充分必要条件是 $g(x)$ 除 $f(x)$ 的余式为零。

【证明】 充分性是显然的。仅证必要性：若 $g(x)\mid f(x)$，则存在 $h(x)\in K[x]$，使得 $f(x)=h(x)g(x)=h(x)g(x)+0$。因 $h(x)$及 0 适合定理 3.1 中对商式 $q(x)$及余式 $r(x)$的要求，由 $q(x)$、$r(x)$ 的唯一性可得 $r(x)=0$。 证毕。

习 题

1. 设 $f_i(x)$、$g_i(x)\in K[x]\,(i=1,2)$，且 $f_1(x)\neq 0$。又若 $f_1(x)f_2(x)$

能被 $g_1(x)g_2(x)$ 整除，$f_1(x)$能整除 $g_1(x)$。求证，$f_2(x)$ 能被 $g_2(x)$整除。

2. 设 $f(x)$、$g_1(x)$、$g_2(x)\in K[x]$，求证

$$(g_1(x)-g_2(x))\mid(f(g_1)-f(g_2)).$$

3. 求 $g(x)$ 除 $f(x)$ 所得的商式及余式：

(i) $f(x)=2x^3+x+1, g(x)=3x^2+x-4$；

(ii) $f(x)=x^4+x^3+2x^2+x-2, g(x)=x^2+3x+1$。

4. m、p、q 适合什么条件时，$g(x)$ 能整除 $f(x)$。

(i) $f(x)=x^3+px+q,\ g(x)=x^2+mx-1$；

(ii) $f(x)=x^4+px^2+q, g(x)=x^2+mx+1$。

§4 最大公因式

一、最大公因式

定义 设 $f(x)$、$g(x)\in K[x]$，称 $d(x)\in K[x]$ 为 $f(x)$ 与 $g(x)$ 的**最大公因式**，如果

(i) $d(x)$ 是 $f(x)$ 与 $g(x)$ 的公因式（即 $d(x)\mid f(x)$，且

$$d(x)\mid g(x)$$）；

(ii) $f(x)$ 与 $g(x)$ 的任一个公因式都能整除 $d(x)$（即只要 $h(x)\mid f(x)$，$h(x)\mid g(x)$，则必有 $h(x)\mid d(x)$）。

公因式的存在性是显然的。因为，每一个非零常数可以作为任意两个多项式的公因式。最大公因式的存在性包含在下面的定理中，其证明基于带余除法。

定理 4.1 $\forall f(x)$、$g(x)\in K[x]$，必存在 $f(x)$ 与 $g(x)$ 的最大公因式 $d(x)\in K[x]$，且有 $u(x)$、$v(x)\in K[x]$，使得

$$u(x)f(x)+v(x)g(x)=d(x). \qquad (1)$$

【证明】 若 $f(x)=g(x)=0$，则 $d(x)=0$，且 $K[x]$ 中任意两个多项式 $u(x)$、$v(x)$ 都可使(1)式成立。若 $f(x)$、$g(x)$ 中有一个为零，不妨设 $f(x)=0$，则 $d(x)=g(x)$，且令 $v(x)=1$，$u(x)$ 为 $K[x]$ 中任一多项式，可使(1)式成立。下面讨论 $f(x)$、$g(x)$ 全不为零的情形。应用带余除法，以 $g(x)$ 除 $f(x)$ 得商式 $q_1(x)$、余式

$r_1(x)$。若 $r_1(x)\neq 0$，则以 $r_1(x)$除 $g(x)$ 得商式 $q_2(x)$、余式 $r_2(x)$。若 $r_2(x)\neq 0$，则再以 $r_2(x)$ 除 $r_1(x)$ 得商式 $q_3(x)$、余式 $r_3(x)$。如此继续下去。因

$$\deg(g)>\deg(r_1)>\deg(r_2)>\deg(r_3)>\cdots\cdots,$$

并且 $\deg(g)$ 为一非负整数，所以这样作了有限次后，必可得一余式 $r_s(x)$，它能整除前一个余式 $r_{s-1}(x)$。即有

$$f(x)=q_1(x)g(x)+r_1(x), \tag{2}$$

$$g(x)=q_2(x)r_1(x)+r_2(x), \tag{3}$$

$$r_1(x)=q_3(x)r_2(x)+r_3(x), \tag{4}$$

$$\cdots\cdots\cdots\cdots$$

$$r_{s-3}(x)=q_{s-1}(x)r_{s-2}(x)+r_{s-1}(x), \tag{5}$$

$$r_{s-2}(x)=q_s(x)r_{s-1}(x)+r_s(x), \tag{6}$$

$$r_{s-1}(x)=q_{s+1}(x)r_s(x)+0, \tag{7}$$

下面说明，$r_s(x)$ 即为 $f(x)$ 与 $g(x)$ 的最大公因式。首先，由 (7) 式，$r_s(x)\mid r_{s-1}(x)$。注意到(6)式并利用整除性质 (iv) 可得

$$r_s(x)\mid r_{s-2}(x).$$

同理 $r_s(x)\mid r_{s-3}(x)\cdots$，如此利用这组等式逐步由下往上推，可得 $r_s(x)\mid g(x)$，$r_s(x)\mid f(x)$，即 $r_s(x)$ 为 $f(x)$ 与 $g(x)$ 的公因式。其次，假设 $h(x)\mid f(x)$，$h(x)\mid g(x)$，注意到 (2) 式并利用整除性质 (iv) 可得 $h(x)\mid r_1(x)$。同理 $h(x)\mid r_2(x)\cdots$，如此利用这组等式逐步由上往下推，最终可得 $h(x)\mid r_s(x)$，即 $r_s(x)$ 为 $f(x)$ 与 $g(x)$的最大公因式。再证明(1)式。由(6)式

$$r_s(x)=r_{s-2}(x)-q_s(x)r_{s-1}(x), \tag{8}$$

再由(5)式，

$$r_{s-1}(x)=r_{s-3}(x)-q_{s-1}(x)r_{s-2}(x). \tag{9}$$

将(9)式代入(8)式并消去$r_{s-1}(x)$，得到

$$r_s(x)=(1+q_s(x)q_{s-1}(x))r_{s-2}(x)-q_s(x)r_{s-3}(x).$$

依此方法，自下而上地利用上一组等式，逐次消去

$$r_{s-2}(x),\ r_{s-3}(x),\ \cdots,\ r_1(x),$$

即可得到

$$r_s(x)=u(x)f(x)+v(x)g(x),$$

其中，$u(x)$、$v(x)$ 为 $K[x]$ 中的多项式。 证毕。

定理 4.1 在证明最大公因式存在性的同时，给出了求最大公因式的一种方法，此方法称为**辗转相除法**。

由最大公因式的定义易知，若 $d_1(x)$、$d_2(x)$ 同为 f 与 g 的最大公因式，则 $d_1(x)|d_2(x)$，且 $d_2(x)|d_1(x)$，于是 $d_1(x)=cd_2(x)$，$c\neq 0$。这表明，两个多项式的最大公因式之间最多相差一个非零常数因子。为确定起见，我们约定，"最大公因式"是指首项系数为 1 的那一个，且当 $f(x)$、$g(x)$ 不全为零时，以记号 (f,g) 来表示 $f(x)$ 与 $g(x)$ 的首项系数为 1 的最大公因式。显然，$(f,g)\neq 0$。今后，称首项系数为 1 的多项式为**首一多项式**。

例 1 设 $f(x)=x^4+x^3-3x^2-4x-1$，$g(x)=x^3+x^2-x-1$，求 (f,g) 及 $u(x)$、$v(x)$ 使得 $u(x)f(x)+v(x)g(x)=(f,g)$。

【解】 按下面格式作辗转相除法。

	$g(x)$	$f(x)$	
	x^3+x^2-x-1	$x^4+x^3-3x^2-4x-1$	$x=q_1(x)$
$q_2(x)=-\frac{x}{2}+\frac{1}{4}$	$x^3+\frac{3x^2}{2}+\frac{x}{2}$	$x^4+x^3-x^2-x$	
	$-\frac{x^2}{2}-\frac{3x}{2}-1$	$r_1(x)=-2x^2-3x-1$	$\frac{8x}{3}+\frac{4}{3}=q_3(x)$
	$-\frac{x^2}{2}-\frac{3x}{4}-\frac{1}{4}$	$-2x^2-2x$	
	$r_2(x)=-\frac{3x}{4}-\frac{3}{4}$	$-x-1$	
		$-x-1$	
		0	

与此相应的一组等式是

$$f(x)=xg(x)+r_1(x),$$

$$g(x)=\left(-\frac{x}{2}+\frac{1}{4}\right)r_1(x)+r_2(x),$$

且 $r_2(x)|r_1(x)$。所以

$$r_2(x)=-\frac{3x}{4}-\frac{3}{4}=g(x)-\left(-\frac{x}{2}+\frac{1}{4}\right)(f(x)-xg(x))$$
$$=\left(\frac{x}{2}-\frac{1}{4}\right)f(x)+\left(-\frac{x^2}{2}+\frac{x}{4}+1\right)g(x),$$

于是得到

$$(f,g)=x+1,$$
$$u(x)=-\frac{2x}{3}+\frac{1}{3},\ v(x)=\frac{2x^2}{3}-\frac{x}{3}-\frac{4}{3}.$$

定理 4.1 中使得(1)式成立的 $u(x)$、$v(x)$ 并不唯一。事实上，对 $f(x)$、$g(x)$ 全为零或其中之一为零的情形，证明中已有叙述。对 $f(x)$、$g(x)$全不为零的情形，若已有$u(x)$、$v(x)$使 (1) 式成立，对任一非零多项式 $h(x)$，令

$$\tilde{u}(x)=u(x)+h(x)g(x),\qquad \tilde{v}(x)=v(x)-h(x)f(x).$$

则 $\tilde{u}(x)\neq u(x)$，$\tilde{v}(x)\neq v(x)$，但 $\tilde{u}(x)$、$\tilde{v}(x)$ 亦能使(1)式成立。此外，仅由等式 $u(x)f(x)+v(x)g(x)=d(x)$，并不能断定 $d(x)$ 就是 f 与 g 的最大公因式。这一点，请初学者务必加以注意。

二、互素的概念及有关性质

定义 设 $f(x)$、$g(x)\in K[x]$。称 $f(x)$ 与 $g(x)$ **互素**，如果

$$(f,g)=1.$$

显然，若两个多项式互素，则它们除去零次多项式外不再有其它的公因式。

定理 4.2 设 $f(x)$、$g(x)\in K[x]$，则 $(f,g)=1$ 的充分必要条件是，存在 $u(x)$、$v(x)\in K[x]$，使得

$$u(x)f(x)+v(x)g(x)=1. \tag{10}$$

【证明】 必要性是显然的。下面证明充分性。设

$$(f,g)=d(x),$$

对(10)式利用整除性质(iv)可得 $d(x)\mid 1$，从而 $d(x)=1$，即

$$(f,g)=1.$$

证毕。

推论 1 若 $f(x)=f_1(x)(f,g)$，$g(x)=g_1(x)(f,g)$，则

$$(f_1, g_1) = 1.$$

事实上,由定理 4.1, 存在 $u(x)$、$v(x)$, 使得

$$u(x)f(x) + v(x)g(x) = (f, g),$$

将 $f(x) = f_1(x)(f, g)$ 及 $g(x) = g_1(x)(f, g)$代入上式并消去(f, g)可得

$$u(x)f_1(x) + v(x)g_1(x) = 1.$$

由定理 4.2 即得$(f_1, g_1) = 1$.

推论 2 若 $(g_1, f) = 1$, $(g_2, f) = 1$, 则 $(g_1g_2, f) = 1$.

证明仅需利用定理 4.2, 留给读者作一练习.

推论 3 若 $f(x) | g_1(x)g_2(x)$, 且 $(f, g_1) = 1$, 则 $f(x) | g_2(x)$.

事实上,由假设 $(f, g_1) = 1$, 有

$$u(x)f(x) + v(x)g_1(x) = 1.$$

以 $g_2(x)$ 乘上式两边得

$$u(x)f(x)g_2(x) + v(x)g_1(x)g_2(x) = g_2(x).$$

再因 $f(x) | g_1(x)g_2(x)$, 故 $f(x)$ 能整除左边两项,从而

$$f(x) | g_2(x).$$

推论 4 若 $f_1(x) | g(x)$, $f_2(x) | g(x)$, 且 $(f_1, f_2) = 1$, 则

$$f_1(x)f_2(x) | g(x).$$

事实上,因 $(f_1, f_2) = 1$, 所以

$$u_1(x)f_1(x) + u_2(x)f_2(x) = 1.$$

以 $g(x)$ 乘上式两边得

$$u_1(x)f_1(x)g(x) + u_2(x)f_2(x)g(x) = g(x).$$

再由整除的假设, 不妨设 $g(x) = q_2(x)f_2(x)$, $g(x) = q_1(x)f_1(x)$, 并分别代入上式左边第一、二项得到

$$u_1(x)q_2(x)f_1(x)f_2(x) + u_2(x)q_1(x)f_1(x)f_2(x) = g(x).$$

显然, $f_1(x)f_2(x)$ 整除左边两项,因而 $f_1(x)f_2(x) | g(x)$.

习　题

1. 求 (f, g) 及 $u(x)$、$v(x)$, 使得

$$u(x)f(x) + v(x)g(x) = (f, g).$$

(i) $f(x)=x^4-x^3-4x^2+4x+1$,

$g(x)=x^2-x-1$;

(ii) $f(x)=x^4-2$,

$g(x)=(x-1)^2$.

2. 设 $d(x)=u(x)f(x)+v(x)g(x)$，且 $d(x)\mid f(x)$，$d(x)\mid g(x)$。求证，$d(x)$是 $f(x)$ 与 $g(x)$ 的最大公因式。

3. 设 $h(x)$ 为首一多项式。证明：

$(f(x)h(x),g(x)h(x))=(f,g)h(x)$。

4. 设 $f(x)$、$g(x)$ 不全为零，且 $u(x)f(x)+v(x)g(x)=(f,g)$。证明：$(u(x),v(x))=1$。

5. 设 $d(x)$ 为首一多项式，$f(x)=f_1(x)d(x)$，$g(x)=g_1(x)d(x)$，且 $(f_1,g_1)=1$。证明：$d(x)=(f,g)$。

6. 设 $f(x)$、$g_1(x)\in K[x]$，且 $f(x)\neq 0$。证明：若对任一适合条件 $f(x)\mid g_1(x)g_2(x)$ 的多项式 $g_2(x)$ 都有 $f(x)\mid g_2(x)$，则 $(f,g_1)=1$。

7. 设 $(f(x),g(x))=1$。求证：$(f(x)g(x),f(x)+g(x))=1$。

8. 设 $f_1(x)$、$f_2(x)\in K[x]$，且全不为零。证明：若对任一适合条件 $f_1(x)\mid g(x)$ 及 $f_2(x)\mid g(x)$ 的多项式 $g(x)$ 都有 $f_1f_2\mid g$，则 $(f_1,f_2)=1$。

9. 证明：如果 $f(x)$ 与 $g(x)$ 互素，则 $f(x^m)$ 与 $g(x^m)$ 也互素，其中 m 为任一正整数。

§5 分解因式定理

在中学代数课程中，将一个多项式分解成因式的乘积（称为“分解因式”），总要进行到“不能再分”时为止。然而，怎样才算“不能再分”呢，当时并未加以追究。本节将讨论一元多项式环中的分解因式问题。首先，须给“不能再分”一个明确的定义。

一、不可约多项式及其性质

定义 设 $p(x)\in K[x]$，且 $\deg(p)\geqslant 1$。称 $p(x)$ 为 K 上的**可约多项式**（或 $p(x)$ 对 K 可约），如果 $p(x)$ 能分解为 K 上两个次数比 $p(x)$ 低的多项式之积。否则，称 $p(x)$ 为 K 上**不可约多项式**（或 $p(x)$ 对 K **不可约**）。

据此定义，零多项式与零次多项式无所谓"可约"与"不可约"。一次式总是不可约的。对于次数大于1的多项式，说它"可约"或"不可约"，必须同时指明其基域。例如，$p(x)=x^2-2$，作为Q上的多项式，它是不可约的；但作为R上的多项式，却是可约的。因为它有且仅有如下的分解式：

$$p(x)=x^2-2=(x+\sqrt{2})(x-\sqrt{2}).$$

上例表明，多项式是否可约依赖于基域，这一点不同于整除，请读者予以注意。

不可约多项式在分解因式中是重要的，下面是有关它的性质。(下文中的"不可约"指对其基域而言)。

命题 1 设 $p(x)\in K[x]$，$\deg(p)\geqslant 1$，则 $p(x)$ 不可约 $\Longleftrightarrow$ $p(x)$只有形如 c 与 $cp(x)$ 的因式.其中 $c\in K$，$c\neq 0$。

【证明】 先证必要性。设 $\phi(x)$ 为 $p(x)$ 的任一个因式，记

$$p(x)=q(x)\phi(x).$$

因 $p(x)$ 不可约，由定义必定有

$$\deg(q)=\deg(p)$$

或者

$$\deg(\phi)=\deg(p),$$

从而只可能是 $\deg(\phi)=0$，$\phi(x)=c(c\neq 0)$或者

$$\phi(x)=cp(x)(c\neq 0).$$

再用反证法证明充分性。若不然，$p(x)$ 可约，由定义存在

$$p_i(x)\in K[x],\ \deg(p_i)<\deg(p),\ (i=1,2),$$

使得

$$p(x)=p_1(x)p_2(x).$$

因 $\deg(p_i)<\deg(p)$，故 $p_i(x)\neq cp(x)$，因而 $p_i(x)\neq c_1(c\neq 0,$ $c_1\neq 0)$。这与假设条件矛盾，从而$p(x)$应为不可约多项式。

证毕。

命题 1 表明，不可约多项式的因式只有两种：非零常数或它自身的一个非零常数倍。

命题 2 若 $p(x)$ 为K 上的不可约多项式，则对 $K[x]$ 中任何

多项式 $f(x)$，或者 $p|f$，或者 $(p, f)=1$。

【证明】 因 $p(x)$ 不可约，由命题 1，其因式只有 c 与 $cp(x)$ 两种($c\neq 0$)。$\forall f(x)\in K[x]$，如果 $cp(x)$ 也是 $f(x)$ 的因式，则 $p|f$；如果 $cp(x)$ 不是 $f(x)$ 的因式，则 $p(x)$ 与 $f(x)$ 的公因式只有非零常数，故 $(p, f)=1$。 证毕。

命题 2 表明，不可约多项式对任一多项式的关系只有两种：或者是它的因式，或者与它互素。

命题 3 若 $p(x)$ 为 K 上的不可约多项式，则对

$$f(x)、g(x)\in K[x],$$

只要 $p|fg$，就必定有 $p|f$ 或者 $p|g$。

【证明】 因 $p(x)$ 不可约，由命题 2，$\forall f(x)\in K[x]$，或者 $p|f$ 或者 $(p, f)=1$。前者 $p|f$ 是结论的一部分；若是后者 $(p, f)=1$，又因 $p|fg$，由定理 4.2 的推论 3，必有 $p|g$。 证毕。

命题 3 表明，一个不可约多项式若能整除两个多项式的乘积，则必定能整除其中之一。利用数学归纳法可将它推广。

命题 3′ 若 $p(x)$ 为 K 上的不可约多项式，则对

$$f_1(x), f_2(x), \cdots, f_s(x)\in K[x], s\geqslant 2,$$

只要 $p|f_1f_2\cdots f_s$，就必定有 $p|f_{i_0}$，其中，$f_{i_0}(x)$ 为

$$f_1(x), f_2(x), \cdots, f_s(x)$$

中的某一个。

二、分解因式定理

下面叙述分解因式的基本定理。

定理 5.1 设 $f(x)\in K[x]$，且 $\deg(f)\geqslant 1$，则

(i) $f(x)$ 必可分解为 K 上的有限个不可约多项式的乘积；

(ii) 如果

$$f(x)=p_1(x)p_2(x)\cdots p_s(x)=q_1(x)q_2(x)\cdots\cdots q_t(x), \tag{1}$$

其中，$p_i(x)$、$q_j(x)(i=1, 2, \cdots, s; j=1, 2, \cdots, t)$ 是 K 上的不可约多项式，那么 $s=t$，且适当调换因式的次序后有

$$p_i(x)=c_iq_i(x), i=1, 2, \cdots, s,$$

其中 $c_i(i=1,2,\cdots,s)$是K上的非零常数。

【证明】 (i) 对 $f(x)$ 的次数作归纳。当 $\deg(f)=1$ 时，注意到一次式总不可约，故结论成立。假设对次数小于 n 的多项式结论成立，当 $\deg(f)=n$ 时，如果 $f(x)$ 不可约，则结论已经成立；如果 $f(x)$ 可约，即有

$$f(x)=f_1(x)f_2(x),$$

其中 $\deg(f_i)<\deg(f)=n(i=1,2)$，由归纳法假设，$f_1(x)$、$f_2(x)$ 可分解为K上的有限个不可约多项式的乘积，不妨记为

$$f_1(x)=p_1(x)p_2(x)\cdots p_l(x),$$
$$f_2(x)=p_{l+1}(x)p_{l+2}(x)\cdots p_s(x),$$

其中 $p_i(x)(i=1,2,\cdots,s)$ 为K上的不可约多项式。将上两式合起来即可得到 $f(x)$ 的分解式

$$f(x)=p_1(x)p_2(x)\cdots p_s(x)。$$

由归纳法原理，结论(i)得证。

(ii) 对 (1) 式中不可约因式的个数 s 作归纳。当 $s=1$ 时，$f(x)=p_1(x)$ 不可约，故必有 $t=1$，$f(x)=p_1(x)=q_1(x)$，结论成立。现假设当分解式中不可约因式的个数为 $s-1$ 时结论成立。由(1) 式，$p_1(x)\mid q_1(x)q_2(x)\cdots q_t(x)$，按命题 3′，$p_1(x)$ 必能整除 $q_1(x)$，$q_2(x)$，$\cdots$，$q_t(x)$ 中之一，适当调换它们的次序，可设 $p_1(x)\mid q_1(x)$。但 $q_1(x)$ 也不可约，所以

$$p_1(x)=c_1q_1(x),\ c_1\neq 0。\tag{2}$$

将(2)式代入(1)式并消去两边的 $q_1(x)$，得到

$$p_2(x)\cdots p_s(x)=c_1^{-1}q_2(x)\cdots q_t(x)。$$

由归纳法假设应有 $s-1=t-1$，即

$$s=t;\tag{3}$$

且适当调换因式的次序后有

$$p_i(x)=c_iq_i(x),\ c_i\neq 0(i=2,3,\cdots,s)。\tag{4}$$

综合(2)、(3)、(4)式，再由归纳法原理即证明了(ii)。 证毕。

定理 5.1 表明，K上任一非常数的多项式必可分解为K上有限个不可约多项式的乘积，且若允许分解式中的因式相差一非零

常数因子，则分解式是唯一的。但是，定理 5.1 并未给出具体的分解方法。实际上，适用于所有多项式的分解因式的普遍方法是不存在的。然而，即便如此，定理 5.1 在多项式的理论中仍有其基本的重要性。

将 $f(x)$ 分解式中的每一个不可约因式的首项系数提出来，使之成为首一多项式，并将相同的不可约因式合并写成幂的形式：

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x)。\tag{5}$$

其中，$c\neq0$ 是 $f(x)$ 的首项系数，$p_i(x)$ 是首一且互异的不可约多项式，r_i 是正整数（$i=1, 2, \cdots, s$）。

称(5)式为 $f(x)$ 的**标准分解式**，它由 $f(x)$ 唯一确定。

三、重因式

定义 称不可约多项式 $p(x)$ 为多项式 $f(x)$ 的 $\boldsymbol{k}$ **重因式**（k 为非负整数），如果 $p^k(x)|f(x)$ 且 $p^{k+1}(x)\nmid f(x)$。

当 $k=0$ 时，$p(x)$ 根本不是 $f(x)$ 的因式；当 $k=1$ 时，称 $p(x)$ 为 $f(x)$ 的**单因式**；当 $k>1$ 时，称 $p(x)$ 为 $f(x)$ 的**重因式**。

容易证明，若 $f(x)$ 的标准分解式为

$$f(x)=ap_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x),$$

则 $p_i(x)$ 是 $f(x)$ 的 r_i 重因式（$i=1, 2, \cdots, s$）。但由于没有一般的方法去求标准分解式，因而判别多项式有无重因式的问题就要另外设法加以解决。为此，引进“导式”的概念。

定义 设 n 次多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$

其（一阶）导式定义为

$$f'(x)=\begin{cases}na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+a_1 & (n\geqslant1)；\\ 0 & (n=0)。\end{cases}$$

由定义，不难验证下列关于导式的公式成立。

(i) $(f(x)+g(x))'=f'(x)+g'(x)$；

(ii) $(f(x)g(x))'=f'(x)g(x)+f(x)g'(x)$；

$(cf(x))'=cf'(x),\ c\neq0$；

(iii) $(f^k(x))' = kf^{k-1}(x)f'(x)$ (k 为正整数)。

称 $f'(x)$ 的导式为 $f(x)$ 的二阶导式，记为 $f''(x)$ 或 $f^{(2)}(x)$；称 $f''(x)$ 的导式为 $f(x)$ 的三阶导式……归纳地，$f(x)$的 k 阶导式定义为

$$f^{(k)}(x) = (f^{(k-1)}(x))',$$

其中，k 为大于 1 的正整数。

显然，一个 n 次多项式的导式是一个$(n-1)$次多项式；它的 n 阶导式是一个非零常数；它的$(n+1)$阶导式是零。

定理 5.2 若不可约多项式 $p(x)$ 是 $f(x)$的 k 重因式($k \geqslant 1$)，则 $p(x)$ 是 $f'(x)$ 的$(k-1)$重因式。

【证明】 由假设

$$f(x) = p^k(x)q(x),$$

其中，$p(x) \nmid q(x)$。对上式求导式得

$$\begin{aligned} f'(x) &= kp^{k-1}(x)p'(x)q(x) + p^k(x)q'(x) \\ &= p^{k-1}(x)(kp'(x)q(x) + p(x)q'(x)), \end{aligned}$$

显然，$p^{k-1}(x) \mid f'(x)$；但 $p(x) \nmid (kp'(x)q(x) + p(x)q'(x))$，否则便有 $p(x) \mid kp'(x)q(x)$，因 $p(x)$ 不可约，且 $p(x) \nmid q(x)$，由命题 3 可得 $p(x) \mid kp'(x)$，但这是不可能的。综上述，$p^{k-1}(x) \mid f'(x)$且

$$p^k(x) \nmid f'(x),$$

即 $p(x)$ 是 $f'(x)$ 的$(k-1)$重因式。 证毕。

定理的特款($k=1$)表明，若不可约多项式 $p(x)$ 是 $f(x)$ 的单因式，则 $p(x)$ 不是 $f'(x)$ 的因式。

下列推论是显然的。

推论 1 不可约多项式 $p(x)$ 是 $f(x)$ 的重因式的充分必要条件是 $p(x)$ 为 $f(x)$ 与 $f'(x)$ 的公因式。

推论 2 $f(x)$ 无重因式的充分必要条件是 $(f(x), f'(x)) = 1$。

推论 2 给出了一种利用多项式的导式判别多项式有无重因式的方法。在 §7 中，还将介绍借助于矩阵的另一种判别方法。

由定理 5.2，还可得到一个去除多项式中因式重数的方法。设 $f(x)$的标准分解式为

$$f(x)=ap_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x),$$

由定理 5.2 可知,

$$(f(x),f'(x))=p_1^{r_1-1}(x)p_2^{r_2-1}(x)\cdots p_s^{r_s-1}(x).$$

令

$$\phi(x)=\frac{f(x)}{(f,f')}=ap_1(x)p_2(x)\cdots\cdots p_s(x),$$

将 $\phi(x)$ 与 $f(x)$ 作一比较可见, $\phi(x)$ 不但具有与 $f(x)$ 完全相同的不可约因式,而且还有次数低及无重因式的优点。在对 $f(x)$ 分解因式或求根时,借助于 $\phi(x)$, 可望简化有关的讨论。

例 试求 $f(x)=x^5-10x^3-20x^2-15x-4$ 在 Q 上的标准分解式。

【解】 $f'(x)=5x^4-30x^2-40x-15$, 用辗转相除法可得

$$(f(x),f'(x))=x^3+3x^2+3x+1.$$

令

$$\phi(x)=\frac{f(x)}{(f,f')}=x^2-3x-4,$$

则 $\phi(x)$ 与 $f(x)$ 具有完全相同的不可约因式。在 Q 上, $\phi(x)$ 有标准分解式

$$\phi(x)=(x-4)(x+1),$$

又易见 $(f,f')=(x+1)^3$(一般情况下,用带余除法可以确定出 (f,f') 中各因式的重数),于是得到 $f(x)$ 在 Q 上的标准分解式为

$$f(x)=(x-4)(x+1)^4.$$

习　题

1. 设数域 K_2 是 K_1 的子域 $(K_2\subseteq K_1)$, 若 $p(x)$ 在 K_1 上不可约, 求证 $p(x)$ 在 K_2 上也不可约。

2. 设 $p(x)\in K[x]$, 且 $\deg(p)>1$。如果对 $K[x]$ 中任何多项式 $f(x)$, 或者 $p|f$, 或者 $(p,f)=1$, 求证, $p(x)$ 在 K 上不可约。

3. 设 $p(x)\in K[x]$, 且 $\deg(p)>1$。如果对所有满足 $p|fg$ 的 $K[x]$ 中的多项式 $f(x)$、$g(x)$, 恒有 $p|f$ 或者 $p|g$. 求证, $p(x)$ 在 K 上不可约。

4. 证明: $x|f^k(x)\Longleftrightarrow x|f(x)$ (k 为正整数)。

5. 证明：$g^2(x)|f^2(x) \Longleftrightarrow g(x)|f(x)$.

6. 判别下列多项式有无重因式，若有，求出它的重因式。

(i) $f(x)=x^4-2x^3-x^2-4x+12$;

(ii) $f(x)=x^3+x+1$.

7. 设 $f(x)=4x^3-4x^2-7x-2$，写出 $f(x)$ 在 Q 上的标准分解式。

§6 多项式函数

定义 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in K[x],\ b\in K,$$

记

$$f(b)=a_nb^n+a_{n-1}b^{n-1}+\cdots+a_1b+a_0, \tag{1}$$

称 $f(b)$ 为 $f(x)$ 在 $x=b$ 处的值。

注意，(1)式右端的运算是 K 中数的运算，其结果 $f(b)$ 自然是 K 中的一个数。

定义 设 $f(x)\in K[x]$，若对 K 中的每一个数 b，都使之与 $f(b)$ 对应，则多项式 $f(x)$ 就定义了 K 上的函数 $f(x)$，称之为**多项式函数**。

特别，当 K 是实数域时，其上的多项式函数就是数学分析中的多项式函数。

按时上述定义，若两个多项式相等，则这两个多项式函数也相等。反之如何？下面的讨论将回答这一问题。

由多项式运算的定义及(1)式，容易得到

命题 1 设 $f(x)$、$g(x)\in K[x]$，若

$$u(x)=f(x)+g(x),\ v(x)=f(x)g(x),$$

则 $\forall b\in K$，恒有

$$u(b)=f(b)+g(b),\ v(b)=f(b)g(b).$$

定理 6.1（余数定理） 设 $f(x)\in K[x]$，$b\in K$，则

$$f(x)=q(x)(x-b)+f(b),$$

其中 $q(x)\in K[x]$。

【证明】 根据带余除法可设

$$f(x)=q(x)(x-b)+c,$$

其中 $q(x)\in K[x]$, $c\in K$。因 $b\in K$, 由命题1应有

$$f(b)=c,$$

即有

$$f(x)=q(x)(x-b)+f(b)。$$

证毕。

推论 $(x-b)|f(x)$的充分必要条件是 $f(b)=0$。

定义 称K中的数b为多项式$f(x)$在K中的**零点或根**，如果$f(b)=0$。

据此定义，上述推论可改述为：b是$f(x)$的根的充分必要条件是$(x-b)$是$f(x)$的因式。于是可进一步定义k重根。

定义 称b为$f(x)$的k**重根**，如果$(x-b)$是$f(x)$的k重因式。当$k=1$时，称b为**单根**，当$k>1$时，称b为**重根**。

定理 6.2 K上任一$n(n\geqslant 1)$次多项式在K中最多有n个不同的根。

【证明】 设$b_1, b_2, \cdots\cdots, b_m$是$f(x)$的$m$个不同的根，由上述推论应有

$$(x-b_j)|f(x) \quad (j=1, 2, \cdots, m)。$$

注意到$x-b_1, x-b_2, \cdots, x-b_m$两两互素，可得

$$f(x)=(x-b_1)(x-b_2)\cdots(x-b_m)q(x),$$

其中，$q(x)$是K上一非零多项式。比较上式两端多项式的次数可得$m\leqslant n$。

证毕。

显然，零多项式在K中有无穷多个(不同的)根。注意到零次多项式无根及定理6.2可知，对数域上的多项式来说，仅零多项式才会有无穷多个(不同的)根。

定理 6.3 设$f(x)$、$g(x)$是K上两个次数不超过n的多项式，若

$$f(b_i)=g(b_i) \quad (i=1, 2, \cdots, n+1),$$

其中$b_1, b_2, \cdots\cdots, b_{n+1}$是$K$中互不相同的数，则

$$f(x)=g(x).$$

【证明】 作多项式 $h(x)=f(x)-g(x)$. 显然，$h(x)$ 的次数不超过 n. 由假设

$$h(b_i)=0,\ (i=1,2,\cdots,n+1),$$

即 $h(x)$ 有 $n+1$ 个不同的根。据定理 6.2，只可能是 $h(x)=0$，即

$$f(x)=g(x).$$ 证毕。

数域 K 中有无穷多个（不同的）数。若 K 上两个多项式函数 $f(x)$ 与 $g(x)$ 相等（即 $\forall b\in K$，恒有 $f(b)=g(b)$），则由定理 6.3 显然可得，这两个多项式 $f(x)$ 与 $g(x)$ 也相等。这就回答了本节开始后不久提出的问题。总之，数域上的多项式，既可当作形式表达式，亦可当作多项式函数。当作形式表达式，便于将多项式的理论用于线性微分方程等其它数学分支，也为定义与讨论一般域上的多项式奠定了基础；而当作多项式函数，则便于运用数学分析的方法加以处理。

设 A 为数域 K 上的 n 阶阵，x 为符号。易知，"形式行列式" $|xI_n+A|$ 是 K 上的 n 次多项式①，且由定理 6.2 可知，K 中有无穷多个 b，使得 bI_n+A 为非异阵。利用这一点，可将对奇异阵的讨论转化为对非异阵的讨论。

例 设 A、B、C、D 为数域 K 上的 n 阶阵，且 $AC=CA$。证明：

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|. \tag{2}$$

【证明】 (i) 若 A 为非异阵，则结论成立（见第二章定理 5.1 推论 2）。

(ii) 若 A 为奇异阵，则在 K 中有无穷多个 b，使得 bI_n+A 为非异阵。注意到，由 $AC=CA$ 可得 $(bI+A)C=C(bI+A)$，再利用(i)的结果，应有

$$\begin{vmatrix} bI_n+A & B \\ C & D \end{vmatrix}=|(bI_n+A)D-CB|. \tag{3}$$

记

① 形式行列式与通常数字方阵的行列式定义相同。下一章将详述。

$$f(x)=\begin{vmatrix} xI_n+A & B \\ C & D \end{vmatrix},\ g(x)=|(xI+A)D-CB|,$$

易知 $f(x)$、$g(x)$ 是次数不超过 n 的多项式。(3) 式表明 $f(x)$ 与 $g(x)$在 K 中无穷多个点上的值相同，据定理 6.3 可得 $f(x)=g(x)$。特别，令 $x=0$，应有 $f(0)=g(0)$，此即

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=|AD-CB|.$$ 证毕。

* 前已述，对于数域，多项式与多项式函数无须加以区别，在那里，看不出形式上定义的多项式较之多项式函数的概念有何拓广之处。然而，对于其它的"域"，情况并不尽然。下面仅就含有两个元素的"有限域" Z_2 加以说明。

在集合 $Z_2=\{\bar{0},\ \bar{1}\}$ 中，规定了如下的"四则"运算：

$$\bar{0}+\bar{0}=\bar{0},\ \bar{0}+\bar{1}=\bar{1},\ \bar{1}+\bar{0}=\bar{1},\ \bar{1}+\bar{1}=\bar{0};$$

$$\bar{0}-\bar{0}=\bar{0},\ \bar{0}-\bar{1}=\bar{1},\ \bar{1}-\bar{0}=\bar{1},\ \bar{1}-\bar{1}=\bar{0};$$

$$\bar{0}\cdot\bar{0}=\bar{0},\ \bar{0}\cdot\bar{1}=\bar{0},\ \bar{1}\cdot\bar{0}=\bar{0},\ \bar{1}\cdot\bar{1}=\bar{1};$$

$$\bar{0}\div\bar{1}=\bar{0},\ \bar{1}\div\bar{1}=\bar{1}.$$

称 Z_2 为含有两个元素的**有限域**，或简称 Z_2 为有限域。

仿照数域 K 上多项式的定义，称形式表达式

$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

为 Z_2 上的多项式，其中 $a_i\in Z_2(i=0,\ 1,\ \cdots,\ n)$，$x$ 为一符号，其它有关的约定及名词都与数域上的多项式完全相同。例如

$$f(x)=\bar{1}x^3+\bar{0}x=x^3,$$

$$g(x)=\bar{1}x^2+x=x^2+x,$$

都是 Z_2 上的多项式。

仿照本章前面各节的讨论，可以证明，有关数域上多项式的一切结论(直至本节定理 6.3 止)，对于 Z_2 上的多项式同样成立。其中，包括"两个多项式相等，则相应的多项式函数必定相等"。但是，值得注意的是，在 Z_2 上，即使两个多项式函数相等，而两个多项式却可以不等。

例如，$f(x)=x^3=\bar{1}x^3$ 与 $g(x)=x^2=\bar{1}x^2$ 是 Z_2 上两个不同的

多项式，然而，由它们定义的 Z_2 上的多项式函数 $f(x)$ 与 $g(x)$ 却是相等的。事实上，$f(\bar{0})=\bar{0}=g(\bar{0})$，$f(\bar{1})=\bar{1}=g(\bar{1})$，即对 Z_2 中任一元素，$f(x)$ 与 $g(x)$ 的值相同。

不难看出，引起上述差别的原因在于，数域中有无限个元素，而有限域 Z_2 中只有两个元素。

在有限域 Z_2 上可以清楚地看到，多项式较之多项式函数的概念更广，它们之间的区别正好说明形式地定义多项式是必要的。

习　题

1．设 $f(x)$ 是数域K上的 n 次多项式，$n\geqslant 1, a\in K$。求证：

$$f(x)=f(a)+f'(a)\frac{x-a}{1!}+f''(a)\frac{(x-a)^2}{2!}+\cdots+f^{(n)}(a)\frac{(x-a)^n}{n!}.$$

2．求证：$1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}$ 不可能有重根。

3．求证：a 是 $f(x)$ 的 k 重根的充分必要条件是

$$f(a)=f'(a)=\cdots=f^{(k-1)}(a)=0, f^{(k)}(a)\neq 0.$$

4．若 a 是 $f'''(x)$ 的一个 k 重根，求证：a 是

$$g(x)=\frac{x-a}{2}[f'(x)+f'(a)]-f(x)+f(a)$$

的一个 $k+3$ 重根。

5．设 $f(x)$ 是 n 次多项式，$n\geqslant 1$，且 $f'(x)|f(x)$。求证：$f(x)$ 有 n 重根。

6．应用二项式定理 $(1+x)^n=c_n^0+c_n^1x+c_n^2x^2+\cdots+c_n^{n-1}x^{n-1}+c_n^nx^n$ 证明 $n(1+x)^{n-1}=c_n^1+2c_n^2x+\cdots+nc_n^nx^{n-1}$ 以及 $\sum\limits_{k=1}^{n}kc_n^k=n\cdot 2^{n-1}$。

7．设 $f(x)$ 是数域K上的多项式，且

$$f(x)=f(x-c),\ c\in K, c\neq 0.$$

证明：$f(x)$ 是零次多项式。

8．设 A、B 为数域K上的同阶非异方阵。求证，存在 $k\in K$，使得 $kA+B$ 与 $kB^{-1}+A^{-1}$ 同为非异阵。

9．设 A 为 n 阶方阵。求证，存在充分小的正数 ε，使得 $I+\varepsilon A$ 的主子式全大于零。

§7 复(实)系数多项式、多项式的友阵

本节讨论的基础是一般数域上多项式的分解因式定理以及代数基本定理。

代数基本定理 任何次数不小于1的复系数多项式必有一个复(数)根。

这一定理由 Gauss 所建立。由于其纯代数的证明冗长且需其它知识，而在“复变函数论”课程中可以简单地证明它，故这里就不证明了。

一、复(实)系数多项式的分解因式定理

利用余数定理的推论，可将代数基本定理改述为：任何次数不小于1的复系数多项式必有一个一次因式 $x-c$(c 为复数)。

由此可知，对于复系数多项式，仅一次式是不可约的，因而，其分解因式定理如下。

定理 7.1 任何 $n(n\geqslant 1)$次复系数多项式 $f(x)$ (在复数域上)必有唯一的标准分解式：

$$f(x)=a_n\prod_{i=1}^{l}(x-c_i)^{r_i},\tag{1}$$

其中 a_n 是 $f(x)$ 的首项系数；$c_i(i=1,2,\cdots,l)$为互异复数；

$$r_i(i=1,2,\cdots,l)$$

是正整数。

如将 k 重根当作 k 个根计算，下面的推论显然成立。

推论 任何 $n(n\geqslant 1)$ 次复系数多项式恰有 n 个复根。

定理 7.2 任何 $n(n\geqslant 1)$ 次实系数多项式(在实数域上)必有唯一的标准分解式：

$$f(x)=a_n\prod_{i=1}^{s}(x-c_i)^{r_i}\cdot\prod_{j=1}^{t}(x^2+p_jx+q_j)^{t_j}.\tag{2}$$

其中，a_n 是 $f(x)$ 的首项系数；$c_i\ (i=1,2,\cdots,s)$ 及 p_j、$q_j\ (j=1,$

$2, \cdots, t$)都是实数；$r_i(i=1, 2, \cdots, s)$ 及 $l_j(j=1,2,\cdots,t)$都是正整数；且 $x^2+p_jx+q_j$ 不可约，即

$$p_j^2-4q_j<0(j=1, 2, \cdots, t).$$

【证明】 设复数 b 是实系数多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 的根，即

$$f(b)=a_nb^n+a_{n-1}b^{n-1}+\cdots+a_1b+a_0=0,$$

考察$f(\bar{b})$，由于

$$\begin{aligned}f(\bar{b})&=a_n\bar{b}^n+a_{n-1}\bar{b}^{n-1}+\cdots+a_1\bar{b}+a_0\\&=\overline{a_nb^n}+\overline{a_{n-1}b^{n-1}}+\cdots+\overline{a_1b}+\overline{a_0}=\bar{0}=0,\end{aligned}$$

所以，$\bar{b}$ 也是 $f(x)$ 的根。且不难证明，作为 $f(x)$ 的根，b 与 $\bar{b}$ 的重数相同。据此性质及(1)式(将 $f(x)$ 看作复系数多项式)，可知 $f(x)$有如下的分解式：

$$f(x)=a_n\prod_{i=1}^{s}(x-c_i)^{r_i}\cdot\prod_{j=1}^{t}(x-b_j)^{l_j}(x-\bar{b}_j)^{l_j}.$$

其中，a_n 是 $f(x)$ 的首项系数；$c_i(i=1, 2, \cdots, s)$ 为互异实数；$r_i(i=1, 2, \cdots, s)$ 及 $l_j(j=1, 2, \cdots, t)$ 都是正整数；$b_j(j=1, 2, \cdots, t)$为复数。又

$$(x-b_j)(x-\bar{b}_j)=x^2-(b_j+\bar{b}_j)x+b_j\bar{b}_j$$

对实数域不可约，记

$$p_j=-(b_j+\bar{b}_j),\ q_j=b_j\bar{b}_j(j=1, 2, \cdots, t),$$

显然 p_j、q_j 为实数，且 $p_j^2-4q_j<0(j=1, 2, \cdots, t)$。于是，(2)式得证。 证毕。

推论 对于实系数多项式，仅一次式以及形如

$$x^2+px+q(p^2-4q<0)$$

的二次式是不可约的。

下面介绍复系数多项式分解因式定理的一个应用。

设 $b_1, b_2, \cdots, b_n$ 是复系数多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

的 n 个复根，$n\geqslant 1$。由定理 7.1，

$$f(x)=a_n(x-b_1)(x-b_2)\cdots(x-b_n).$$

比较上式两端同次项的系数，可得

$$b_1+b_2+\cdots+b_n=\sum_{i=1}^{n}b_i=-\frac{a_{n-1}}{a_n},$$

$$b_1b_2+b_1b_3+\cdots+b_1b_n+b_2b_3+\cdots+b_{n-1}b_n$$

$$=\sum_{1\le i_1<i_2\le n}b_{i_1}b_{i_2}=\frac{a_{n-2}}{a_n},$$

$$\sum_{1\le i_1<i_2<\cdots<i_k\le n}b_{i_1}b_{i_2}\cdots b_{i_k}=(-1)^k\frac{a_{n-k}}{a_n},\tag{3}$$

………………………………………………………………

$$b_1b_2\cdots b_n=(-1)^n\frac{a_0}{a_n}.$$

这一组式子给出了复系数多项式根与系数之间的关系，称为 **Vieta 公式**。

二、两个多项式的公根问题

两个多项式有无公根的问题，在几何及工程技术问题中都会遇到，这里介绍一种判别方法。先引进多项式的友阵的概念。

由于任一复系数多项式都可等于一个数乘上一个首一多项式，为方便起见，在讨论有关根的问题时，总假设所讨论的多项式是首一多项式。

定义 设

$$f(x)=x^n+a_1x^{n-1}+a_2x^{n-2}+\cdots+a_{n-1}x+a_n\tag{4}$$

是数域 K 上的多项式(函数)，$n\geqslant 1$，作 n 阶 Frobenius 阵：

$$F=\begin{pmatrix}0 & & & & -a_n\\ 1 & 0 & & & -a_{n-1}\\ & 1 & \ddots & & \vdots\\ & & \ddots & 0 & -a_2\\ & & & 1 & -a_1\end{pmatrix},\tag{5}$$

称 F 为 $f(x)$ 的**友阵**。

如 $f(x)=x^3-2x^2+3$ 的友阵是

$$\begin{pmatrix} 0 & 0 & -3 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

由行列式的计算可知,

$$|xI_n - F| = f(x). \tag{6}$$

称$|xI_n - F|$为 F 的(关于变量 x 的)**特征多项式**. (6)式说明,任何多项式(函数)可以看作它的友阵的特征多项式.

定义 设 $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ 是数域 K 上的多项式, $n \geqslant 1$. A 为 K 上的方阵,记

$$f(A) = a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I_s,$$

上式中加法、乘法与数乘分别是矩阵的加法、乘法与数乘. 称 $f(A)$ 为 $f(x)$ 在 $x = A$ 时的(矩阵)**值**.

由 $f(A)$ 的定义,显然可得到如下的命题.

命题 1 若数域 K 上的多项式 $f(x)$、$g(x)$、$h(x)$、$l(x)$ 使得

$$h(x) = f(x) + g(x),\ l(x) = f(x)g(x),$$

则对 K 上的任一方阵 A, 必有

$$h(A) = f(A) + g(A),\ l(A) = f(A)g(A).$$

定理 7.3 (Barnett) 设 $f(x)$、$g(x)$ 是首一的复系数多项式, $\deg(f) = n \geqslant 1$, $\deg(g) = m \geqslant 1$, F 是 $f(x)$ 的友阵,则

$$(f, g) = 1 \Longleftrightarrow g(F) \text{非异}.$$

【证明】 由条件,不妨设 $g(x)$ 的标准分解式为

$$g(x) = \prod_{j=1}^{s} (x - \zeta_j)^{r_j},$$

根据命题 1,

$$g(F) = \prod_{j=1}^{s} (F - \zeta_j I_n)^{r_j}.$$

两边取行列式,得到

$$|g(F)| = \prod_{j=1}^{s} |F - \zeta_j I_n|^{r_j}.$$

注意到 F 是 $f(x)$ 的友阵,由(6)式可知,

$$|F-\zeta_j I_n| = (-1)^n |\zeta_j I_n - F| = (-1)^n f(\zeta_j),$$
$$(j=1, 2, \cdots, s).$$

于是得到

$$|g(F)| \neq 0 \Longleftrightarrow f(\zeta_j) \neq 0 \qquad (j=1, 2, \cdots, s).$$

而 $f(\zeta_j)\neq 0 (j=1,2,\cdots,s)$表明 $f(x)$与 $g(x)$无公根，亦即$(f,g)=1$。

证毕。

下列推论是显而易见的。

推论 1 设 $f(x)$、$g(x)$ 是首一且次数不小于 1 的复系数多项式，F 为 $f(x)$ 的友阵。则 $f(x)$ 与 $g(x)$ 有公根 $\Longleftrightarrow g(F)$ 为奇异阵。

推论 2 设 $f(x)$ 是首一且次数不小于 1 的复系数多项式，F 为 $f(x)$ 的友阵。则 $f(x)$ 有重根 $\Longleftrightarrow f'(F)$ 为奇异阵。

由于在定理 7.3 中 $f(x)$ 与 $g(x)$ 的地位是对等的，故若互易它们的位置，相应的结论自然也成立。

在利用上述结论去判断多项式有无公根(或重根)时，主要的一步是计算友阵的幂。下面介绍当幂指数 k 不超过友阵 F 的阶数 n 时，求 F^k 的一种简易可行的方法——先求出 F^k 的第 1 列，再由第 1 列推出第 2 列，由第 2 列推出第 3 列……，直至推出第 n 列。

将 F^k 按列分块为 $F^k=(\beta_1, \beta_2, \cdots, \beta_n)$，不难证明

$$\beta_j = F\beta_{j-1} \qquad (j=2, 3, \cdots, n). \tag{7}$$

事实上，因 F 是 Frobenius 阵，故对 $j=2, 3, \cdots, n$，有

$$e_j = Fe_{j-1},$$

于是，

$$\beta_j = F^k e_j = F^k(Fe_{j-1}) = F(F^k e_{j-1}) = F\beta_{j-1}.$$

(7)式给出了 F^k 中后一列与前一列的关系。

又

$$\beta_1 = F^k e_1 = F^{k-1}(Fe_1) = F^{k-1}e_2 = \cdots$$
$$= \begin{cases} Fe_k = e_{k+1}, & \text{当 } k<n \text{ 时}; \\ Fe_n = \alpha, & \text{当 } k=n \text{ 时}, \end{cases} \tag{8}$$

其中，α 是 F 的末列向量。(8)式给出了求 F^k 的第1列的方法。

例1 判断 $f(x)=x^5+x^4+1$ 与 $g(x)=x^2-x+4$ 有无公共复根。

【解】 注意到 $\deg(f)>\deg(g)$，为了使得在计算友阵的幂时不出现幂指数超过阶数的情形，我们作 $f(x)$ 的友阵

$$F=\begin{pmatrix}0&0&0&0&-1\\1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&-1\end{pmatrix},$$

并考察 $g(F)$ 的非异性。由(7)、(8)式易得

$$g(F)=F^2-F+4I_5$$

$$=\begin{pmatrix}0&0&0&-1&1\\0&0&0&0&-1\\1&0&0&0&0\\0&1&0&0&0\\0&0&1&-1&1\end{pmatrix}+\begin{pmatrix}0&0&0&0&1\\-1&0&0&0&0\\0&-1&0&0&0\\0&0&-1&0&0\\0&0&0&-1&1\end{pmatrix}$$

$$+\begin{pmatrix}4&0&0&0&0\\0&4&0&0&0\\0&0&4&0&0\\0&0&0&4&0\\0&0&0&0&4\end{pmatrix}=\begin{pmatrix}4&0&0&-1&2\\-1&4&0&0&-1\\1&-1&4&0&0\\0&1&-1&4&0\\0&0&1&-2&6\end{pmatrix};$$

注意到 $g(F)$ 是严格对角占优阵，故 $|g(F)|\neq 0$，所以 $f(x)$ 与 $g(x)$ 没有公根。

对上例，如若考虑 $g(x)$ 的友阵

$$F_1=\begin{pmatrix}0&-4\\1&1\end{pmatrix},$$

算得

$$f(F_1)=F_1^5+F_1^4+I_2=\begin{pmatrix}41&8\\-2&39\end{pmatrix}.$$

因 $f(F_1)$ 非异，也能得到 $f(x)$ 与 $g(x)$ 没有公根的结论。这里，要计算 F_1 的高次幂 F_1^k（k 大于 F_1 的阶数），当 F_1 的阶数大时，较为麻烦。但在下一章将看到，F_1^k 可化为 F_1^l，而 l 不超过 F_1 的阶数。

例 2 解方程组

$$\begin{cases} y^2-7xy+4x^2+13x-2y-3=0 \\ y^2-14xy+9x^2+28x-4y-5=0. \end{cases} \tag{9}$$

【解】 将(9)改写为

$$\begin{cases} y^2-(7x+2)y+(4x^2+13x-3)=0 \\ y^2-2(7x+2)y+(9x^2+28x-5)=0. \end{cases} \tag{10}$$

视方程左端为关于 y 的一元多项式，记

$$f_x(y)=y^2-(7x+2)y+(4x^2+13x-3),$$

$$g_x(y)=y^2-2(7x+2)y+(9x^2+28x-5).$$

则显然，(x_0, y_0) 为方程组的解 $\Longleftrightarrow y_0$ 是 $f_{x_0}(y)$ 与 $g_{x_0}(y)$ 的公根。

注意到 $f_x(y)$ 的友阵为

$$F=\begin{pmatrix} 0 & -(4x^2+13x-3) \\ 1 & 7x+2 \end{pmatrix},$$

由 Barnett 定理的推论 1 可知，$f_x(y)$ 与 $g_x(y)$ 有公根的充要条件是 $|g_x(F)|=0$，即

$$|F^2-2(7x+2)F+(9x^2+28x-5)I_2|=0.$$

经计算，上式为

$$-\begin{vmatrix} 5x^2+15x-2 & 28x^3+99x^2+5x-6 \\ 7x+2 & 44x^2+13x+6 \end{vmatrix}=0,$$

亦即有

$$-24x(x-1)(x+2)(x-2)=0,$$

故其根 x_0 可为 $0, 1, -2, 2$。将 $x_0=0$ 代入(10)，得到

$$\begin{cases} y^2-2y-3=0 \\ y^2-4y-5=0, \end{cases}$$

解得 $y_0=-1$，故$(0, -1)$为(9)的一个解。用同样的方法可得(9)的另三个解是$(1, 2)$，$(-2, 1)$，$(2, 3)$。于是，方程组(9)共有四

个解：

$$(0, -1), (1, 2), (-2, 1), (2, 3).$$

习　题

1. 设 3 次多项式 $f(x)=x^3+px^2+qx+r$ 有一根是其余两根之和，求证：$p^3-4pq+8r=0$。

2. 设 $(x^2+x+1)\mid(f_1(x^3)-xf_2(x^3))$，求证：

$$(x-1)\mid f_1(x), (x-1)\mid f_2(x).$$

3. 设 3 次多项式 $f(x)=x^3+px^2+qx+r$ 的三个根成等差数列，求证：$2p^3-9pq+27r=0$。反之如何？

4. 用定理 7.3 判断

(i) $f(x)$ 与 $g(x)$ 是否有公根：

$$f(x)=x^4+2x^3+3x^2+x-11, g(x)=x^3+x+1;$$

(ii) $f(x)=6x^4+x^2-1$ 是否有重根。

5. 证明：$x^3+px+q=0$ 有重根 $\Longleftrightarrow 4p^3+27q^2=0$。

6. 用定理 7.3 找出 $f(x)$ 与 $g(x)$ 有公根的充要条件：

$$f(x)=x^3+px+q,\quad g(x)=x^4+2px+p.$$

7. 找数 p，使 $f(x)=x^n-px-1$ 与 $g(x)=x^n+x+p$ 没有公根。其中，$n\geqslant 2$。

8. 证明：对任何有理数 b、c 以及 $a\neq 0$，必有

$$4a^3+2b^3+c^3-6abc\neq 0.$$

（提示：注意到 x^3-2 与 $x^2+\frac{b}{a}x+\frac{c}{a}$ 没有公根，并应用 Barnett 定理。）

§8　有理系数多项式

有理系数多项式的分解因式虽然有不少应用，然而比起复系数与实系数多项式的分解因式却要复杂得多。因为，复系数多项式的不可约因式都是一次式；实系数多项式的不可约因式最多是二次式；然而，却存在任意次的有理系数多项式对有理数域不可约。本节就来证明这一点。

对任一有理系数多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$，

只要乘上 $f(x)$ 的各个系数 a_i 的分母的最小公倍数 c，则 $cf(x)$就是一个整系数多项式。

定义 称整系数多项式 $f(x)$ 为**本原多项式**，如果 $f(x)$ 的各个系数的最大公因数是 1。

引理 1 两个本原多项式的乘积仍是本原多项式。

【证明】 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0$$

是两个本原多项式，如果

$$f(x)g(x)=c_{n+m}x^{n+m}+c_{n+m-1}x^{n+m-1}+\cdots+c_1x+c_0$$

不是本原多项式，则 $c_0, c_1, \cdots, c_{n+m}$ 的最大公因数 $d\neq 1$。设 p 是 d 的一个素因子。因为 $f(x)$ 是本原的，所以 p 不能同时整除 $f(x)$ 的所有的系数，不妨设

$$p|a_0,\ p|a_1,\ \cdots,\ p|a_{i-1},\ p\nmid a_i;$$

又因 $g(x)$ 也是本原的，同理可设

$$p|b_0,\ p|b_1,\ \cdots,\ p|b_{j-1},\ p\nmid b_j.$$

而

$$c_{i+j}=\cdots+a_{i-2}b_{j+2}+a_{i-1}b_{j+1}+a_ib_j+a_{i+1}b_{j-1}$$
$$+a_{i+2}b_{j-2}+\cdots,$$

上式右端除 a_ib_j 以外的各项都能被 p 整除，所以 $p|a_ib_j$，于是 $p|a_i$ 或 $p|b_j$，这与假设相矛盾，故 $f(x)g(x)$ 是本原多项式。

证毕。

引理 2 设非零的整系数多项式 $f(x)$ 在有理数域 Q 上可约，则 $f(x)$ 必能分解为两个次数较 $f(x)$ 低的整系数多项式的乘积。

【证明】 由假设

$$f(x)=g(x)h(x), \tag{1}$$

其中，$g(x)$、$h(x)$ 都是 Q 上的多项式，且 $g(x)$、$h(x)$ 的次数都小于 $f(x)$ 的次数。设 c 是 $g(x)$ 的各个系数的分母的最小公倍数，记 $cg(x)=g_1(x)$，则 $g_1(x)$ 是整系数多项式。把 $g_1(x)$ 的各系数的最大公因数 d 提出，写成 $g_1(x)=d\phi(x)$，则 $\phi(x)$ 显然是本原多项

式，于是，

$$g(x)=\frac{d}{c}\phi(x)=s\phi(x), \tag{2}$$

其中，s 是有理数。同理，可将 $h(x)$ 写成如下形式：

$$h(x)=t\psi(x), \tag{3}$$

其中，t 是有理数，$\psi(x)$ 是本原多项式。将(2)、(3)式代入(1)式，得到

$$f(x)=st\phi(x)\psi(x)=\frac{m}{n}\phi(x)\psi(x), \tag{4}$$

其中，m 与 n 是互素的整数。

因为 $f(x)$ 是整系数多项式，所以 $\phi(x)\psi(x)$ 的每一个系数 c_k 乘 m 后必须能被 n 整除。但 m 与 n 互素，所以每一个 c_k 必能被 n 整除。又由引理 1，$\phi(x)\psi(x)$ 也是本原多项式，所以 $n=\pm1$，于是(4)式化为

$$f(x)=(\pm m\phi(x))\psi(x),$$

其中，$\pm m\phi(x)$ 与 $\psi(x)$ 都是整系数多项式，且次数都低于 $f(x)$ 的次数。　　证毕。

定理 8.1(Eisenstein)　设整系数多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \qquad (a_n\neq0,\ n\geqslant1),$$

如果存在素数 p，使得 $p|a_i(i=0,1,2,\cdots,n-1)$，$p\nmid a_n$，且 $p^2\nmid a_0$，则 $f(x)$ 在有理数域上不可约。

【证明】　若不然，即 $f(x)$ 在 Q 上可约，则由引理 2，$f(x)$ 有分解式

$$f(x)=(b_lx^l+\cdots+b_0)(c_mx^m+\cdots+c_0),$$

其中，b_i、$c_j(i=0,1,\cdots,l;j=0,1,\cdots,m)$都是整数，且 l、$m<n$，$l+m=n$。易知

$$a_n=b_lc_m,\ a_0=b_0c_0.$$

由假设 $p|a_0$，所以 $p|b_0$ 或者 $p|c_0$；但因 $p^2\nmid a_0$，故 p 不能同时整除 b_0 与 c_0，不妨设 $p|b_0$，但 $p\nmid c_0$。

另一方面，由假设 $p\nmid a_n$，于是 $p\nmid b_l$，进而可以断定：必存在 $k(0<k\leqslant l)$，使得 $p|b_j(j=0,1,\cdots,k-1)$，但 $p\nmid b_k$。考察 $f(x)$

中 x^k 的系数

$$a_k = b_k c_0 + b_{k-1} c_1 + \cdots + b_0 c_k,$$

因 $k \leqslant l < n$，由假设条件，有 $p|a_k$。又上式右端除 $b_k c_0$ 以外的各项都能被 p 整除，所以也应有 $p|b_k c_0$，而这与 $p \nmid c_0$、$p \nmid b_k$ 相矛盾。故 $f(x)$ 在 Q 上必不可约。

证毕。

由定理 8.1，可构造出任意次有理系数多项式，在有理数域上不可约。例如

$$f(x) = x^n + px + p,$$

其中，n 为任一正整数，p 是素数。正因为如此，故讨论有理系数多项式的分解因式相当复杂，至今还没有什么好的解决方法。

习　题

1. 设 p 是素数，求证：$x^n - p$ 在有理数域上不可约。

2. 证明：

(i) 如果有理系数多项式 $f(x)$ 在 Q 上不可约，则 $f(x+1)$ 在 Q 上也不可约；

(ii) 若 p 为素数，则 $x^{p-1} + x^{p-2} + \cdots + x + 1$ 在 Q 上不可约。

3. 求证 $f(x) = x^4 - 2x + 3$ 在 Q 上不可约。

4. 求证，$px^2 - px + (2p-1)$ 在 Q 上不可约（p 是素数）。

5. 设 $p_1, p_2, \cdots, p_r$ 是互异的素数，$m > 1$。求证，$\sqrt[m]{p_1 p_2 \cdots p_r}$ 必定不是有理数。

6. 设 $a_1, a_2, \cdots, a_n$ 是互异的整数。求证，

$$f(x) = \prod_{i=1}^{n} (x - a_i) - 1$$

在 Q 上不可约。

§9　多元多项式

为了实用上的需要及处理问题简化起见，以下均设 $x_1, x_2, \cdots, x_n$ 是取自某一数域 K 的 n 个变数。称 $ax_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ 为一个单项式。其中，$a \in K$ 称为此单项式的系数；k_i 是非负整数，当 $a \neq 0$ 时，称

为此单项式中 x_i 的次数($i=1, 2, \cdots, n$),而称 $k=k_1+k_2+\cdots+k_n$ 为此单项式的次数。

如果两个单项式中的所有 x_i 的次数对应相同,则称这两个单项式是**同类项**。同类项的次数必定相同,但反之不然(见下例(1)式),这种情形在一元多项式中是不会出现的。

一些单项式的和

$$f(x_1, x_2, \cdots, x_n)=\sum a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n},$$

称为 **n 元多项式**(函数),在不致混淆的情况下,简称为多项式,也记为 $f(x)$。今后,我们总认为 n 元多项式中的各个单项式彼此不同类,也就是同类项已进行了"合并"(即所有同类项前的系数已相加)。

在 n 元多项式中,称系数非零的单项式的最大次数为这个多项式的次数。例如,三元多项式

$$f(x_1, x_2, x_3)=x_1^3x_2x_3-x_2x_3^2-3x_1^2x_2x_3^2+x_3^5 \tag{1}$$

的次数是5。

如果 n 元多项式中不存在系数非零的单项式,则称它为**零多项式**。

定义 称 n 元多项式 $f(x)$ 与 $g(x)$ **相等**,如果它们同类项的系数全部对应相等。

命题1 数域 K 上的 n 元多项式 $f(x_1, x_2, \cdots, x_n)$ 与 $g(x_1, x_2, \cdots, x_n)$ 相等的充分必要条件是,对于 K 中的任一组数 $\xi_1, \xi_2, \cdots, \xi_n$,恒有

$$f(\xi_1, \xi_2, \cdots, \xi_n)=g(\xi_1, \xi_2, \cdots, \xi_n).$$

【证明】 必要性是显然的,下面证充分性(对 n 作归纳)。

当 $n=1$ 时,f 与 g 为一元多项式,由§6的讨论可知,命题成立;

今假设命题对 $n-1$ 元多项式成立。由于 $n-1\geqslant 1$,可设

$$f(x_1, x_2, \cdots, x_n)=\Sigma a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$
$$=\sum_{i=0}^{r} a_i(x_1, x_2, \cdots, x_{n-1})x_n^i, \tag{2}$$

$$g(x_1, x_2, \cdots, x_n) = \sum b_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$$
$$= \sum_{j=0}^{s} b_j(x_1, x_2, \cdots, x_{n-1}) x_n^j, \tag{3}$$

其中，$a_i(x_1, x_2, \cdots, x_{n-1})$ 及 $b_j(x_1, x_2, \cdots, x_{n-1})$ 为元数不超过 $n-1$ 的多项式 ($i=0, 1, \cdots, r$; $j=0, 1, \cdots, s$)，且 $a_r(x_1, x_2, \cdots, x_{n-1}) \neq 0$，$b_s(x_1, x_2, \cdots, x_{n-1}) \neq 0$。不难证明 $r=s$。事实上，因 $a_r(x_1, x_2, \cdots, x_{n-1}) \neq 0$，故总可在 K 中取到一组(固定的)数 $x_1^0, x_2^0, \cdots, x_{n-1}^0$，使 $a_r(x_1^0, x_2^0, \cdots, x_{n-1}^0) \neq 0$。由命题的假设条件，$\forall x_n \in K$，应有

$$f(x_1^0, x_2^0, \cdots, x_{n-1}^0, x_n) = g(x_1^0, x_2^0, \cdots, x_{n-1}^0, x_n),$$

即

$$a_r(x_1^0, x_2^0, \cdots, x_{n-1}^0) x_n^r + \cdots = b_s(x_1^0, x_2^0, \cdots, x_{n-1}^0) x_n^s + \cdots.$$

视上式为关于 x_n 的两个一元多项式相等，即可知 $r>s$ 是不可能的。相仿地，可证 $r<s$ 也不可能。于是(2)与(3)式可写成

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=0}^{r} a_i(x_1, x_2, \cdots, x_{n-1}) x_n^i, \tag{4}$$

$$g(x_1, x_2, \cdots, x_n) = \sum_{j=0}^{r} b_j(x_1, x_2, \cdots, x_{n-1}) x_n^j. \tag{5}$$

由于取 K 中任意一组数 $x_1=\xi_1$，$x_2=\xi_2$，$\cdots$，$x_{n-1}=\xi_{n-1}$ 时，(4)与(5)式都是(关于 x_n 的)一元多项式，类似于"$r=s$"的证明可知，$\forall \xi_1, \xi_2, \cdots, \xi_{n-1} \in K$，恒有

$$a_i(\xi_1, \xi_2, \cdots, \xi_{n-1}) = b_i(\xi_1, \xi_2, \cdots, \xi_{n-1})$$
$$(i=0, 1, \cdots, r).$$

由归纳假设，可以断定 $a_i(x_1, x_2, \cdots, x_{n-1})$ 与 $b_i(x_1, x_2, \cdots, x_{n-1})$ 相等，即 $a_i(x_1, x_2, \cdots, x_{n-1})$ 与 $b_i(x_1, x_2, \cdots, x_{n-1})$ 的同类项的系数全都对应相等。再注意到(4)、(5)两式，即知 $f(x_1, x_2, \cdots, x_n)$ 与 $g(x_1, x_2, \cdots, x_n)$ 的同类项的系数也全都对应相等。

证毕。

两个 n 元多项式相加，即是将它们的同类项的系数相加。例如

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_1 x_3,$$

$$g(x_1, x_2, x_3) = 4x_2^2 - 7x_1x_2 - 3x_1x_3,$$

则

$$f(x_1, x_2, x_3) + g(x_1, x_2, x_3) = x_1^2 + 5x_2^2 - 7x_1x_2 - 2x_1x_3.$$

两个 n 元多项式相乘，即是将一多项式的各项遍乘另一多项式的各项，并且合并其中的同类项。例如

$$f(x_1, x_2) = 3x_1^2 - x_1x_2 + x_2^2,$$

$$g(x_1, x_2) = x_1 - x_2,$$

则
$$f(x_1, x_2)g(x_1, x_2) = 3x_1^3 - 4x_1^2x_2 + 2x_1x_2^2 - x_2^3.$$

前已述，在多元多项式中，次数相同的单项式可以不同类，故无法对其中的单项式按次数给出一个自然排列顺序，这给许多问题的讨论带来了不便。下面引入**字典排列法**，它能给多元多项式中的单项式以唯一确定的排列顺序。

以 $x_1, x_2, \cdots, x_n$ 作为自然顺序。对于两个不同的单项式

$$ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}, \tag{6}$$

$$bx_1^{l_1}x_2^{l_2}\cdots x_n^{l_n}, \tag{7}$$

如果 $k_1 = l_1, k_2 = l_2, \cdots, k_{s-1} = l_{s-1}$，而 $k_s > l_s (s \geqslant 1)$，则称 (6) 先于(7)或(7)后于(6)。例如，单项式 $x_1x_2^3x_3^2x_4$ 先于 $2x_1x_2^3x_3x_4^5x_5$。易知，只要两个单项式不同类，其中总有一个先于另一个。

单项式之间的这样一种先后关系具有传递性。设另有一单项式

$$cx_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}, \tag{8}$$

若(6)先于(7)，(7)先于(8)，则不难证明，(6)必先于(8)。

将多项式中的各项，按上述规定的先后顺序排列，称为**字典排列法**。例如，多项式

$$f(x) = 5x_1^4x_2 + 2x_1^2x_2^3x_4 - x_1^2x_2^2x_3^3 + x_1^2x_2^2x_3x_4 + x_3^7$$

就是按照字典排列法排列的一个 4 元 7 次多项式。

按照字典排列法写出的多项式的第 1 项，称为多项式的**首项**。由上例可见，首项的次数未必就是多项式的次数。

命题 2 设 $f(x_1, x_2, \cdots, x_n) \neq 0, g(x_1, x_2, \cdots, x_n) \neq 0$，则乘积 $f(x_1, x_2, \cdots, x_n)g(x_1, x_2, \cdots, x_n)$ 的首项等于 f 的首项与 g 的首

项的乘积。

【证明】 设 $ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 与 $bx_1^{l_1}x_2^{l_2}\cdots x_n^{l_n}$ 分别是 f 与 g 的首项，则 $a\neq0$, $b\neq0$。又设 $cx_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}$ 与 $dx_1^{r_1}x_2^{r_2}\cdots x_n^{r_n}$ 分别是 f 与 g 的任一其它单项式，则 $f(x_1, x_2, \cdots, x_n)g(x_1, x_2, \cdots, x_n)$中的单项式无非是下列四种类型：

$$abx_1^{k_1+l_1}x_2^{k_2+l_2}\cdots x_n^{k_n+l_n}, \tag{9}$$

$$adx_1^{k_1+r_1}x_2^{k_2+r_2}\cdots x_n^{k_n+r_n}, \tag{10}$$

$$cbx_1^{m_1+l_1}x_2^{m_2+l_2}\cdots x_n^{m_n+l_n}, \tag{11}$$

$$cdx_1^{m_1+r_1}x_2^{m_2+r_2}\cdots x_n^{m_n+r_n}. \tag{12}$$

由于 $ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 先于 $cx_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}$，故(9)先于(11)；又由于 $bx_1^{l_1}x_2^{l_2}\cdots x_n^{l_n}$ 先于 $dx_1^{r_1}x_2^{r_2}\cdots x_n^{r_n}$，故(9)先于(10)，且(11)先于(12)。再由传递性可知，(9)也先于(12)。于是(9)先于(10)、(11)、(12)，即(9)是乘积 fg 的首项。 证毕。

由命题 2 显然可得下列推论。

推论 1 如果 $f(x_1, x_2, \cdots, x_n)\neq0$, $g(x_1, x_2, \cdots, x_n)\neq0$，则 $f(x_1, x_2, \cdots, x_n)g(x_1, x_2, \cdots, x_n)\neq0$。

推论 2 如果

$$f(x_1, x_2, \cdots, x_n)h(x_1, x_2, \cdots, x_n)$$
$$=g(x_1, x_2, \cdots, x_n)h(x_1, x_2, \cdots, x_n),$$

且 $h(x_1, x_2, \cdots, x_n)\neq0$，则

$$f(x_1, x_2, \cdots, x_n)=g(x_1, x_2, \cdots, x_n).$$

下面介绍齐次多项式的概念。

定义 称 $f(x_1, x_2, \cdots, x_n)$是 **s 次齐次多项式**或 **s 次型**，如果 $f(x_1, x_2, \cdots, x_n)$的所有单项式都是 s 次的。

例如，$\sum_{i=1}^{n}a_ix_i$ 称为**一次型**或**线性型**；$\sum_{i,j=1}^{n}a_{ij}x_ix_j$ 称为**二次型**；又如 $3x_1^4+x_1^2x_2^2$ 是 4 **次型**。

显然，两个齐次多项式的乘积仍是齐次多项式，它的次数就等于这两个多项式次数的和。

任何一个 n 元多项式都可唯一地表示为若干个齐次多项式之

出。事实上，这只要把所有次数相同的单项式集合在一起就能得出所需要的表示式。例如

$$\begin{aligned}f(x_1, x_2, x_3, x_4) &= 4x_1^2x_4^2 + x_1x_2 - x_1x_3^2 + x_2^2x_3 + 2x_3^2 + x_4^4 + 6\\ &= (4x_1^2x_4^2 + x_4^4) + (-x_1x_3^2 + x_2^2x_3)\\ &\quad + (x_1x_2 + 2x_3^2) + 6.\end{aligned}$$

这里，$f(x_1, x_2, x_3, x_4)$ 是 4 次型、3 次型、2 次型与零次型之和。

命题 3 设 $f(x_1, x_2, \cdots, x_n) \neq 0$，$g(x_1, x_2, \cdots, x_n) \neq 0$，则乘积 $f(x_1, x_2, \cdots, x_n)g(x_1, x_2, \cdots, x_n)$ 的次数等于 f 的次数与 g 的次数之和。

【证明】 设 f 与 g 的次数分别为 s 与 t。先将 f 与 g 表示为齐次多项式之和：

$$f(x_1, x_2, \cdots, x_n) = \phi(x_1, x_2, \cdots, x_n) + \cdots, \tag{13}$$

$$g(x_1, x_2, \cdots, x_n) = \psi(x_1, x_2, \cdots, x_n) + \cdots, \tag{14}$$

其中，$\phi(x_1, x_2, \cdots, x_n)$为 s 次型，(13)式中未写出的部分都是次数比 s 小的齐次型；$\psi(x_1, x_2, \cdots, x_n)$为 t 次型，(14)式中未写出的部分都是次数比 t 小的齐次型。因为

$$\begin{aligned}&f(x_1, x_2, \cdots, x_n)g(x_1, x_2, \cdots, x_n)\\ &= \phi(x_1, x_2, \cdots, x_n)\psi(x_1, x_2, \cdots, x_n) + \cdots,\end{aligned}$$

其右端第一项为 $(s+t)$ 次型，其余部分都是次数比 $s+t$ 小的齐次型，所以 $f(x_1, x_2, \cdots, x_n)g(x_1, x_2, \cdots, x_n)$的次数为 $s+t$。

证毕。

§10 对称多项式

在多元多项式中，最有用的是对称多项式。

定义 称 $f(x_1, x_2, \cdots, x_n)$ 为(关于 $x_1, x_2, \cdots, x_n$ 的)**对称多项式**，如果对任何 $i \neq j$，$1 \leqslant i, j \leqslant n$，恒有

$$f(x_1, \cdots, \overset{i}{x_i}, \cdots, \overset{j}{x_j}, \cdots, x_n) = f(x_1, \cdots, \overset{i}{x_j}, \cdots, \overset{j}{x_i}, \cdots, x_n).$$

例如，$x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2$ 是齐次对称多项式；$x_1x_2 + x_1x_3 + x_2x_3 + x_1x_2x_3$ 是非齐次对称多项式。

在 n 元对称多项式中，下列 n 个多项式是最基本的：

$$\sigma_1 = x_1 + x_2 + \cdots + x_n = \sum_{i=1}^{n} x_i,$$

$$\sigma_2 = x_1x_2 + x_1x_3 + \cdots + x_1x_n + x_2x_1 + \cdots + x_{n-1}x_n$$

$$= \sum_{1\leqslant i<j\leqslant n} x_ix_j,$$

$$\cdots\cdots\cdots\cdots$$

$$\sigma_n = x_1x_2\cdots x_n.$$

称 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 为**初等(或基本)对称多项式**。

设 $g(y_1, y_2, \cdots, y_n)$ 是任一多项式，记

$$g(\sigma_1, \sigma_2, \cdots, \sigma_n) \equiv f(x_1, x_2, \cdots, x_n).$$

则由于每一个 σ_i 都是(关于 $x_1, x_2, \cdots, x_n$ 的)对称多项式，又对称多项式的和与积仍是对称多项式，所以 $f(x_1, x_2, \cdots, x_n)$ 也是对称多项式。这一结论的逆也是正确的。

定理 10.1 对于数域 K 上的任一对称多项式 $f(x_1, x_2, \cdots, x_n)$，必可找到 K 上唯一的一个多项式 $g(y_1, y_2, \cdots, y_n)$，使得

$$f(x_1, x_2, \cdots, x_n) = g(\sigma_1, \sigma_2, \cdots, \sigma_n). \tag{1}$$

【证明】 先证存在性。设 $f(x_1, x_2, \cdots, x_n)$（按字典排列法）的首项为

$$ax_1^{l_1}\cdots x_i^{l_i}\cdots x_k^{l_k}\cdots x_n^{l_n},\ a\neq 0. \tag{2}$$

由于 f 是对称多项式，故必有

$$l_1 \geqslant l_2 \geqslant \cdots \geqslant l_n. \tag{3}$$

事实上，若不然，有 $l_i < l_k (i < k)$，则由 f 的对称性可知，f 中必有另一项

$$ax_1^{l_1}\cdots x_i^{l_k}\cdots x_k^{l_i}\cdots x_n,$$

但这一项先于(2)，从而与(2)为首项的假设相矛盾，所以(3)式成立。

由(3)式，作关于 $x_1, x_2, \cdots, x_n$ 的多项式：

$$a\sigma_1^{l_1-l_2}\sigma_2^{l_2-l_3}\cdots\sigma_{n-1}^{l_{n-1}-l_n}\sigma_n^{l_n} \equiv g_1(x_1, x_2, \cdots, x_n). \tag{4}$$

显然，g_1 是关于 $x_1, x_2, \cdots, x_n$ 的对称多项式，且易知 g_1 的首项恰好是

$$ax_1^{l_1-l_2}(x_1x_2)^{l_2-l_3}\cdots(x_1x_2\cdots x_n)^{l_n}=ax_1^{l_1}x_2^{l_2}\cdots x_n^{l_n},$$

即与 f 的首项相同，所以多项式

$$f_1=f-g_1$$

的首项必然后于 f 的首项，且易知 f_1 也是关于 $x_1, x_2, \cdots, x_n$ 的对称多项式，其系数仍在 K 中。对 f_1 重复上述做法，并且继续下去，就得到了 K 上的一串对称多项式：

$$f_0=f,\ f_1=f_0-g_1,\ f_2=f_1-g_2,\ \cdots$$

且 f_i 的首项后于 f_{i-1} 的首项($i=1, 2, \cdots$)。

设 $bx_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 是任一 $f_i(i>1)$ 的首项，则因它后于 f 的首项，故有

$$l_1\geqslant k_1\geqslant k_2\geqslant\cdots\geqslant k_n\geqslant 0.$$

满足上述条件的 $k_1, k_2, \cdots, k_n$ 只能有有限个，所以多项式 f_i 不能无限地构造下去，即存在正整数 s，使得 $f_s=0$，从而

$$f=f_0=g_1+f_1=g_1+g_2+f_2=\cdots=g_1+g_2+\cdots+g_s.$$

由于每一个 g_i 都形如(4)，故可将 $f(x_1, x_2, \cdots, x_n)$ 表示为初等对称多项式的一个多项式：

$$f=g(\sigma_1, \sigma_2, \cdots, \sigma_n),$$

$g(y_1, y_2, \cdots, y_n)$ 即为所求的使(1)式成立的 K 上的一个多项式。

再证唯一性。设 $g(y_1, y_2, \cdots, y_n)$ 及 $h(y_1, y_2, \cdots, y_n)$ 都是 K 上的多项式，使得

$$\begin{aligned}f(x_1, x_2, \cdots, x_n)&=g(\sigma_1, \sigma_2, \cdots, \sigma_n)\\&=h(\sigma_1, \sigma_2, \cdots, \sigma_n).\end{aligned}\tag{5}$$

令

$$\phi(y_1, y_2, \cdots, y_n)=g(y_1, y_2, \cdots, y_n)-h(y_1, y_2, \cdots, y_n),$$

则(5)式即为

$$\phi(\sigma_1, \sigma_2, \cdots, \sigma_n)=0.\tag{6}$$

若 $\phi(y_1, y_2, \cdots, y_n)$ 不是零多项式，不妨设

$$\begin{aligned}\phi(y_1, y_2, \cdots, y_n)&=ay_1^{k_1}y_2^{k_2}\cdots y_n^{k_n}+by_1^{l_1}y_2^{l_2}\cdots y_n^{l_n}\\&\quad+\cdots+cy_1^{r_1}y_2^{r_2}\cdots y_n^{r_n},\end{aligned}\tag{7}$$

其中，$a, b, \cdots, c$ 全不为零，且所有的单项式彼此不同类。考察

$\phi(\sigma_1, \sigma_2, \cdots, \sigma_n)$，由

$$a\sigma_1^{k_1}\sigma_2^{k_2}\cdots\sigma_n^{k_n} = ax_1^{k_1}(x_1x_2)^{k_2}\cdots(x_1x_2\cdots x_n)^{k_n} + \cdots$$
$$= ax_1^{k_1+k_2+\cdots+k_n}x_2^{k_2+k_3+\cdots+k_n}\cdots x_{n-1}^{k_{n-1}+k_n}x_n^{k_n} + \cdots;$$

以及

$$b\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n} = bx_1^{l_1+l_2+\cdots+l_n}x_2^{l_2+l_3+\cdots+l_n}\cdots x_{n-1}^{l_{n-1}+l_n}x_n^{l_n} + \cdots;$$
$$c\sigma_1^{r_1}\sigma_2^{r_2}\cdots\sigma_n^{r_n} = cx_1^{r_1+r_2+\cdots+r_n}x_2^{r_2+r_3+\cdots+r_n}\cdots x_{n-1}^{r_{n-1}+r_n}x_n^{r_n} + \cdots.$$

并注意到(7)式中无同类项，可知

$$a\sigma_1^{k_1}\sigma_2^{k_2}\cdots\sigma_n^{k_n},\ b\sigma_1^{l_1}\sigma_2^{l_2}\cdots\sigma_n^{l_n}\ \text{以及}\ c\sigma_1^{r_1}\sigma_2^{r_2}\cdots\sigma_n^{r_n}$$

的首项互不相同，因此

$$\phi(\sigma_1, \sigma_2, \cdots, \sigma_n) \neq 0.$$

这与(6)式相矛盾，故 $\phi(y_1, y_2, \cdots, y_n)$ 只能是零多项式，亦即

$$g(y_1, y_2, \cdots, y_n) = h(y_1, y_2, \cdots, y_n).$$

证毕。

定理 10.1 也称为**对称多项式的基本定理**，其存在性的证明是构造性的，可用来将对称多项式表示为初等对称多项式的一个多项式。

例 1 设 $n \geqslant 3$，试将对称多项式

$$\Sigma x_1^2x_2 \equiv x_1^2x_2 + x_1^2x_3 + \cdots + x_1^2x_n + x_1x_2^2 + \cdots + x_1x_n^2 + x_2^2x_3$$
$$+ x_2^2x_4 + \cdots + x_2^2x_n + \cdots + x_{n-1}x_n^2 \text{①}$$

表示成 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式。

【解】 用定理 10.1 证明的方法。对首项 $x_1^2x_2$，作

$$\sigma_1^{2-1}\sigma_2^1 = \sigma_1\sigma_2 = (x_1 + x_2 + \cdots + x_n)(x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n)$$
$$\equiv (\Sigma x_1)(\Sigma x_1x_2) = \Sigma x_1^2x_2 + 3\Sigma x_1x_2x_3,$$

故

$$\Sigma x_1^2x_2 = \sigma_1\sigma_2 - 3\Sigma x_1x_2x_3 = \sigma_1\sigma_2 - 3\sigma_3.$$

当 n 较大时，上述方法常常是麻烦的，而采用“待定系数法”却比较方便。下面以齐次对称多项式为例加以说明。

例 2 将对称多项式

---

① $\Sigma x_1^2x_2$ 代表由形如 $x_i^2x_j$ ($i \neq j$, $1 \leqslant i, j \leqslant n$) 的项所构成的 n 元 3 次齐次对称多项式。

$$f(x_1, x_2, \cdots, x_n) = \Sigma x_1^2 x_2^2$$

表示成 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式。

【解】 因为 f 是 n 元 4 次齐次对称多项式，故若它的首项为 $x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n}$，则

$$l_1 + l_2 + \cdots + l_n = 4,$$

且

$$l_1 \geqslant l_2 \geqslant \cdots \geqslant l_n \geqslant 0.$$

(i) 当 $n=2$ 时，$f(x_1, x_2) = x_1^2 x_2^2 = \sigma_2^2$；

(ii) 当 $n=3$ 时，由定理 10.1，$f(x_1, x_2, x_3)$ 表示成 $\sigma_1, \sigma_2, \sigma_3$ 的多项式的整个构造过程可如下表所示：

可能出现的首项中 x_1, x_2, x_3 的次数排列	可能的首项	相应的 g_i
220	$x_1^2 x_2^2$	$\sigma_1^{2-2} \sigma_2^2 = \sigma_2^2$
211	$A x_1^2 x_2 x_3$	$A\sigma_1^{2-1} \sigma_2^{1-1} \sigma_3^1 = A\sigma_1\sigma_3$

其中 A 为待定的系数。且此时有 $f = g_1 + g_2$，即

$$x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 = \sigma_2^2 + A\sigma_1\sigma_3.$$

在上式两边取 $x_1 = x_2 = x_3 = 1$，解得 $A = -2$。所以

$$x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 = \sigma_2^2 - 2\sigma_1\sigma_3.$$

(iii) 当 $n \geqslant 4$ 时，同 (ii) 一样，$f(x_1, x_2, \cdots, x_n)$ 表示成 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式的构造过程如下表所示：

可能出现的首项中 $x_1, x_2, \cdots, x_n$ 的次数排列	相应的 g_i
2 2 0 0 ⋯ 0	$\sigma_1^{2-2} \sigma_2^2 = \sigma_2^2$
2 1 1 0 ⋯ 0	$A\sigma_1^{2-1} \sigma_2^{1-1} \sigma_3^1 = A\sigma_1\sigma_3$
1 1 1 1 ⋯ 0	$B\sigma_1^{1-1} \sigma_2^{1-1} \sigma_3^{1-1} \sigma_4 = B\sigma_4$

其中 A、B 均为待定的系数。于是有

$$\Sigma x_1^2 x_2^2 = \sigma_2^2 + A\sigma_1\sigma_3 + B\sigma_4.$$

在上式两边取 $x_1 = x_2 = x_3 = 1$，$x_4 = x_5 = \cdots = x_n = 0$，解得 $A = -2$；

再取 $x_1=x_2=x_3=x_4=1$, $x_5=x_6=\cdots=x_n=0$, 又得 $B=2$。所以

$$\Sigma x_1^2x_2^2=\sigma_2^2-2\sigma_1\sigma_2+2\sigma_4.$$

当所给多项式非齐次时,可先将它表示成若干个齐次型的和,然后再用待定系数法。

定理 10.1 在讨论多项式的根与系数的关系时,亦经常被用到。

例 3 设 $c_1, c_2, \cdots, c_n$ 是数域K上 n 次多项式 $h(x)$ 的 n 个复根($n\geqslant 1$)。

(i) 求证,对K上任一个 n 元对称多项式 $f(x_1,x_2,\cdots,x_n)$, 恒有

$$f(c_1, c_2, \cdots, c_n)\in K;$$

(ii) 若 $h(x)=x^n+a_1x^{n-1}+\cdots+a_n$, 试以 $a_1, a_2, \cdots, a_n$ 表示出 $h(x)$ 的所有复根的平方和。

【解】

(i) 因 $f(x_1, x_2, \cdots, x_n)$ 对称, 由定理 10.1, f 可唯一地表示成初等对称多项式的多项式:

$$f(x_1, x_2, \cdots, x_n)=g(\sigma_1, \sigma_2, \cdots, \sigma_n).$$

从而

$$f(c_1, c_2, \cdots, c_n)=g\left(\sum_{i=1}^n c_i, \sum_{1\leqslant i<j\leqslant n} c_ic_j, \cdots, \prod_{i=1}^n c_i\right).$$

又由 Vieta 公式可知

$$\sum_{i=1}^n c_i, \sum_{1\leqslant i<j\leqslant n} c_ic_j, \cdots, \prod_{i=1}^n c_i$$

全属于 K, 所以

$$f(c_1, c_2, \cdots, c_n)\in K.$$

(ii) 特别令 $f(x_1, x_2, \cdots, x_n)=\sum_{i=1}^n x_i^2$, 容易得到

$$\sum_{i=1}^n x_i^2=\sigma_1^2-2\sigma_2.$$

从而

$$\sum_{i=1}^n c_i^2=\left(\sum_{i=1}^n c_i\right)^2-2\left(\sum_{1\leqslant i<j\leqslant n} c_ic_j\right).$$

再利用 Vieta 公式，可知

$$\sum_{i=1}^{n} c_i^2 = (-a_1)^2 - 2a_2 = a_1^2 - 2a_2.$$

习　题

1. 将下列对称多项式表示为初等对称多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式：

(i) $\Sigma x_1^2 x_2 x_3$;

(ii) $\Sigma x_1^2 x_2^2 x_3 x_4$;

(iii) $\Sigma x_1^3 x_2$;

(iv) $\Sigma x_1^3 x_2 x_3$.

2. 已知 $x^4 + ax^3 + bx^2 + cx + d = 0$ 有两根的积等于另两根的积，试求系数 a、b、c、d 之间的关系。

3. 已知 $x^3 + px^2 + qx + r = 0$ 的根是 c_1, c_2, c_3，试求一个三次多项式 $f(x) = x^3 + ax^2 + bx^2 + c$，使得 $f(x)$ 的根恰好是 c_1^2, c_2^2, c_3^2.

选 做 题

1. 设 $f(x)$、$\phi(x) \in K[x]$，且 $\deg(\phi) > 0$. 证明：存在 $c_i(x) \in K[x]$ $(i = 0, 1, \cdots, k)$，使得

$$f(x) = c_k(x)\phi^k(x) + c_{k-1}(x)\phi^{k-1}(x) + \cdots + c_1(x)\phi(x) + c_0(x).$$

其中 $c_i(x) = 0$ 或者 $\deg(c_i) < \deg(\phi)$；且这样的 $c_i(x)$ $(i = 0, 1, \cdots, k)$ 由 $f(x)$ 与 $\phi(x)$ 唯一确定。

2. 证明：若 $\dfrac{f(x)}{(f,g)}$ 及 $\dfrac{g(x)}{(f,g)}$ 的次数都大于零，则必存在适合条件

$$\deg(u) < \deg\left(\frac{g}{(f,g)}\right), \deg(v) < \deg\left(\frac{f}{(f,g)}\right)$$

的多项式 $u(x)$、$v(x)$，使得

$$u(x)f(x) + v(x)g(x) = (f,g);$$

且这样的 $u(x)$、$v(x)$ 由 $f(x)$ 与 $g(x)$ 唯一确定。

3. 设 $\deg(f) \geqslant 1$. 证明：$f(x)$ 等于某个不可约多项式 $p(x)$ 的 k $(k \geqslant 1)$ 次幂的充分必要条件是，对任何多项式 $g(x)$，必有 $(f,g) = 1$ 或者

$$f(x) \mid g^m(x)$$

(m 为某一正整数).

4. 设 $a_1, a_2, \cdots, a_n$ 是互不相同的数，

$$f(x)=\prod_{i=1}^{n}(x-a_i),\quad g_i(x)=\frac{f(x)}{(x-a_i)f'(a_i)}\quad(i=1,2,\cdots,n).$$

求证

(i) $g_i(a_j)=\delta_{ij}(i,j=1,2,\cdots,n)$;

(ii) $\sum_{i=1}^{n} g_i(x)=1$;

(iii) 对任意一组数 $b_1,b_2,\cdots,b_n$, 必存在多项式

$$L(x)=\sum_{i=1}^{n} b_i g_i(x),$$

使得

$$L(a_j)=b_j\quad(j=1,2,\cdots,n).$$

称这个 $L(x)$ 为 **Lagrange 内插多项式**.

5. 设 $A=\overline{P}'[\lambda_1,\lambda_2,\cdots,\lambda_n]P$, 其中 P 为 n 阶酉阵 (即 $\overline{P}'P=P\overline{P}'=I_n$), $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是一组复数. 求证, 必存在多项式 $f(x)$、$g(x)$, 使得

$$\overline{A}'=f(A),\quad A=g(\overline{A}').$$

6. 设 $a_1,a_2,\cdots,a_n$ 是 n 次首一多项式 $f(x)$ 的根.

(i) 记

$$V(a_1,a_2,\cdots,a_n)=\prod_{1\leqslant i<j\leqslant n}(a_j-a_i),$$

求证,

$$V^2(a_1,a_2,\cdots,a_n)=(-1)^{\frac{n(n-1)}{2}}f'(a_1)f'(a_2)\cdots f'(a_n);$$

(ii) 记

$$s_k=\sum_{i=1}^{n} a_i^k\quad(k\text{ 为非负整数}).$$

求证,

$$\begin{vmatrix} s_0 & s_1 & \cdots & s_{n-1}\\ s_1 & s_2 & \cdots & s_n\\ \cdots & \cdots & \cdots & \cdots\\ s_{n-1} & s_n & \cdots & s_{2n-2}\end{vmatrix}=(-1)^{\frac{n(n-1)}{2}}f'(a_1)f'(a_2)\cdots f'(a_n).$$

7. 记 $s_k=\sum_{i=1}^{n}x_i^k$, $k=0,1,2,\cdots$. 证明下列牛顿(Newton)公式:

$$s_k-s_{k-1}\sigma_1+s_{k-2}\sigma_2-s_{k-3}\sigma_3+\cdots+(-1)^{k-1}s_1\sigma_{k-1}+(-1)^k k\sigma_k=0$$

(当 $k\leqslant n$ 时);

$$s_k-s_{k-1}\sigma_1+s_{k-2}\sigma_2-s_{k-3}\sigma_3+\cdots+(-1)^n s_{k-n}\sigma_n=0$$

(当 $k>n$ 时).

第五章　方阵的特征值、特征多项式与最小多项式

本章介绍在理论上与工程技术上都有广泛应用的方阵的特征值的概念，讨论它的一些应用。引进与特征值相联系的方阵的特征多项式的概念，介绍著名的 Hamilton-Cayley 定理以及与此定理相关联的方阵的最小多项式的概念。

§1　特征值与特征向量

一、特征值的概念

方阵的特征值，最早是一个物理概念，是 Laplace 在十九世纪为研究天体力学、地球力学等而引进的。这一概念，不仅在理论上极为重要，在技术科学领域里，它的应用也是多种多样的。事实上，在讨论振动问题(如机械振动、弹性体振动、电磁波震荡)、天体运行问题及现代控制理论中，都涉及到特征值问题。下面从数学上引进这一概念。

例1　求二次齐次函数 $\sum\limits_{i,j=1}^{n} a_{ij}x_ix_j$ 在"超球面" $\sum\limits_{i=1}^{n} x_i^2=1$ 上的(限制)极值。其中，$a_{ij}=a_{ji}(i,\ j=1,\ 2,\ \cdots,\ n)$。

【解】　用 Lagrange 乘数法。考虑多元函数

$$F(x_1,\ x_2,\ \cdots,\ x_n)=\sum_{i,j=1}^{n} a_{ij}x_ix_j-\lambda\left(\sum_{i=1}^{n} x_i^2-1\right),$$

如果$(x_1,\ x_2,\ \cdots,\ x_n)'$ 是极值点，则

$$\frac{\partial F}{\partial \lambda}=0,\quad \frac{\partial F}{\partial x_i}=0\quad (i=1,\ 2,\ \cdots,\ n).$$

由 $\frac{\partial F}{\partial \lambda}=0$ 可导出原限制条件 $\sum_{i=1}^{n} x_i^2=1$，而由 $\frac{\partial F}{\partial x_i}=0 \quad (i=1, 2, \cdots, n)$ 容易得到

$$\sum_{j=1}^{n} a_{ij}x_j-\lambda x_i=0 \qquad (i=1, 2, \cdots, n). \tag{1}$$

如果记 $A=(a_{ij})_{n\times n}$，则(1)式可以写成

$$Ax=\lambda x, \tag{2}$$

亦即

$$(\lambda I_n - A)x=0. \tag{3}$$

其中，$x=(x_1, x_2, \cdots, x_n)'$，并且 $x \neq 0$（因为 $x'x=1$）。

这样，寻找 $F(x_1, x_2, \cdots, x_n)$的极值点的问题，化为寻找方程组(2)或(3)的非零解的问题。能使方程组(2)或(3)有非零解的数 λ 以及相关的非零解 x，就是下面要引进的方阵 A 的特征值与特征向量。

从现在起，如果矩阵 A 的元素均取自某一数域 K，则称 A 是数域 K 上的矩阵，称 K 为基域。同样，如果 n 维列(行)向量 x 的分量均取自 K，则称 x 是 K 上的 n 维列(行)向量。

定义 设 $A=(a_{ij})_{n\times n}$ 是数域 K 上的 n 阶阵。如果存在数 $\lambda\in K$ 以及 K 上的 n 维非零列向量 x，使得

$$(\lambda I_n - A)x=0(\text{或 } Ax=\lambda x),$$

则称 λ 是 A 的特征值(或特征根)，而称 x 是属于 λ 的特征向量(有时也称 λ 是相应于特征向量 x 的特征值)。

称 $\lambda I_n - A$ 为 A 的特征矩阵。称 $|\lambda I_n - A|$ 为 A 的特征多项式，以 $f_A(\lambda)$记之。由行列式的定义可知，$f_A(\lambda)$ 是次数为 n 的首一多项式。称 $|\lambda I_n - A|=0$ 为 A 的特征方程(有时也称 $|\lambda I_n - A|=0$ 为 A 的永年方程或长期方程，这是由于天文学上的原因)。

根据方程组的理论，容易知道，数域 K 上的方阵 A 的特征值就是其特征多项式在 K 上的根；A 的属于特征值 λ_0 的特征向量就是齐次方程组$(\lambda_0 I-A)x=0$ 的所有的非零解，它们有无限多个。今后，称齐次方程组 $(\lambda_0 I-A)x=0$ 的任一个基础解系为 A 的属于特征值 λ_0 的极大线性无关特征向量组。

例 2 求 3 阶阵

$$A=\begin{pmatrix}1 & -1 & 1\\ 1 & 3 & -1\\ 1 & 1 & 1\end{pmatrix}$$

的所有的特征值，并求出属于每一个特征值的极大线性无关特征向量组。

【解】 因为A的特征多项式

$$|\lambda I_3-A|=\begin{vmatrix}\lambda-1 & 1 & -1\\ -1 & \lambda-3 & 1\\ -1 & -1 & \lambda-1\end{vmatrix}=(\lambda-1)(\lambda-2)^2.$$

所以A的特征值是：$\lambda_1=1,\ \lambda_2=\lambda_3=2$。

对 $\lambda_1=1$，齐次线性方程组

$$(\lambda_1 I_3-A)x=\begin{pmatrix}0 & 1 & -1\\ -1 & -2 & 1\\ -1 & -1 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

的一个基础解系是

$$\begin{pmatrix}-1\\ 1\\ 1\end{pmatrix}.$$

所以，属于 $\lambda_1=1$ 的极大线性无关特征向量组为$(-1,\ 1,\ 1)'$。

对 $\lambda_2=2$，齐次线性方程组

$$(\lambda_2 I_3-A)x=\begin{pmatrix}1 & 1 & -1\\ -1 & -1 & 1\\ -1 & -1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

即 $x_1+x_2-x_3=0$ 的一个基础解系是

$$\begin{pmatrix}1\\ 0\\ 1\end{pmatrix},\ \begin{pmatrix}0\\ 1\\ 1\end{pmatrix}.$$

所以，属于 $\lambda_2=2$ 的极大线性无关特征向量组为 $(1,\ 0,\ 1)'$ 与$(0,\ 1,\ 1)'$。

例 3 求 3 阶阵

$$A=\begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 2 \end{pmatrix}$$

的所有的特征值以及属于每一个特征值的极大线性无关特征向量组.

【解】 与例 2 一样,容易算出 $|\lambda I_3 - A| = (\lambda-2)^3$, 而属于$\lambda=2$ 的极大线性无关特征向量组是$(1, -1, 0)'$ 与$(1, -1, 1)'$.

二、特征值与特征向量的性质

并非数域K上的任何方阵(在K上)都有特征值,例如,实方阵

$$A=\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

在实数域上无特征值. 但是, 若将$|\lambda I_n - A|$看作复数域上的多项式, 则由代数基本定理可知它必有 n 个复根. 于是关于特征值的存在性有如下的结论.

命题 1 数域K上的任何 n 阶阵必有 n 个复的特征值.

特别地,对于 n 阶(实)对称阵,可以证明它的 n 个特征值全是实数.

定理 1.1 对称阵的特征值全是实数.

【证明】 将对称阵A看作复方阵, 设 λ 是 A的任一特征值, x 是属于 λ 的特征向量,则

$$Ax=\lambda x,\ x\neq 0. \tag{4}$$

对上式两边转置共轭,得

$$\bar{x}'\bar{A}'=\bar{\lambda}\bar{x}', \tag{5}$$

由假设 A是(实)对称阵,即 $\bar{A}'=A'=A$, 所以(5)式可写成

$$\bar{x}'A=\bar{\lambda}\bar{x}', \tag{6}$$

(4) 式的两边左乘 $\bar{x}'$, (6)式的两边右乘 x, 即得

$$\bar{x}'Ax=\lambda\bar{x}'x,$$
$$\bar{x}'Ax=\bar{\lambda}\bar{x}'x,$$

所以

$$(\lambda-\bar{\lambda})\bar{x}'x=0.$$

由于 $x\neq0$，故 $\bar{x}'x\neq0$，于是 $\lambda=\bar{\lambda}$，即 λ 是一实数。　　证毕。

虽然定理 1.1 的证明比较简单，然而在历史上，这个定理的证明曾被认为是困难的。当时，这个结论的提出(作为猜测)，甚至使一些有名的数学家感到"惊异"；而企图论证这一定理所用的方法也是错误的。例如，Laplace 运用了时间的无限性这一物理概念来论证这一数学命题；甚至提出这一命题的 Cauchy(1829 年)，其证明也是不完全的。

定理 1.2　设 λ_1，λ_2，…，λ_s 是 n 阶阵 A 的 s 个不同的特征值，α_1，α_2，…，α_s 分别是属于 λ_1，λ_2，…，λ_s 的特征向量，则 α_1，α_2，…，α_s 必线性无关。

【证明】 对 s 作归纳。当 $s=1$ 时，定理显然正确，因为一个非零向量必线性无关。现假设有数 k_1，k_2，…，k_s，使得

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0, \tag{7}$$

上式两边左乘 A，并利用 $A\alpha_i=\lambda_i\alpha_i(i=1, 2, \cdots, s)$可得

$$k_1\lambda_1\alpha_1+k_2\lambda_2\alpha_2+\cdots+k_s\lambda_s\alpha_s=0. \tag{8}$$

又(7)式乘以 λ_s，再减去(8)式，得到

$$k_1(\lambda_s-\lambda_1)\alpha_1+k_2(\lambda_s-\lambda_2)\alpha_2+\cdots+k_{s-1}(\lambda_s-\lambda_{s-1})\alpha_{s-1}=0.$$

由归纳法假设应有

$$k_i(\lambda_s-\lambda_i)=0 \quad (i=1, 2, \cdots, s-1),$$

但因 $\lambda_s-\lambda_i\neq0$，所以 $k_i=0(i=1, 2, \cdots, s-1)$。将它代入(7)式，即得 $k_s\alpha_s=0$。但 $\alpha_s\neq0$，所以 $k_s=0$，于是 α_1，α_2，…，α_s 线性无关。　　证毕。

由定理 1.2 易知，相应于一个特征向量的特征值只有一个。

定理 1.3　设 λ_1，λ_2，…，λ_s 是 n 阶阵 A 的 s 个不同的特征值，α_{i1}，α_{i2}，…，α_{ir_i} 是属于 λ_i 的线性无关的特征向量 $(i=1, 2, \cdots, s)$，则 α_{11}，α_{12}，…，α_{1r_1}，α_{21}，…，α_{sr_s} 必线性无关。

【证明】 设有

$$\sum_{j=1}^{r_1} k_{1j}\alpha_{1j} + \sum_{j=1}^{r_2} k_{2j}\alpha_{2j} + \cdots + \sum_{j=1}^{r_s} k_{sj}\alpha_{sj} = 0,$$

若记 $\xi_i = \sum_{j=1}^{r_i} k_{ij}\alpha_{ij} (i = 1, 2, \cdots, s)$，则上式即为

$$\xi_1 + \xi_2 + \cdots + \xi_s = 0. \tag{9}$$

由此可知 $\xi_1, \xi_2, \cdots, \xi_s$ 全为零。否则，其中不为零的 ξ_i 是属于 λ^i 的特征向量，由(9)式便得到属于不同特征值的特征向量线性相关而与定理 1.2 矛盾。

由 $$\sum_{j=1}^{r_i} k_{ij}\alpha_{ij} = 0 \qquad (i = 1, 2, \cdots, s),$$

又注意到 $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{ir_i}$ 线性无关，故 $k_{i1} = k_{i2} = \cdots = k_{ir_i} = 0 (i = 1, 2, \cdots, s)$，这就证明了 $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1r_1}, \alpha_{21}, \cdots, \alpha_{sr_s}$ 线性无关。 证毕。

推论 设 n 阶阵 A 有 s 个不同的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$，$A$ 的属于 λ_i 的极大线性无关特征向量组中向量的个数为 $n_i (i = 1, 2, \cdots, s)$，则

$$n_1 + n_2 + \cdots + n_s \leqslant n,$$

且等号成立的充要条件是 A 有 n 个线性无关的特征向量。（充分性的证明留作习题）

定义 如果 n 阶阵 A 有 n 个线性无关的特征向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$，则称 A 有完全特征向量系，同时称 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 A 的完全特征向量系。

上面例 2 的 A 在有理数域上有完全特征向量系。例 3 的 A，即便在复数域上也没有。

习 题

1．求下列方阵的特征值及属于这些特征值的极大线性无关特征向量组。

(i) $\begin{pmatrix} 1-\sqrt{-1} & 1 & \sqrt{-1} \\ 0 & \sqrt{-1} & 0 \\ 0 & 0 & -2 \end{pmatrix}$； (ii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$；

(iii) $\begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$; (iv) $\begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ -1 & -2 & 0 \end{pmatrix}$.

2．设λ是n阶阵A的特征值，α是A的属于λ的特征向量．求证：

(i) $k\lambda$是kA的特征值，且α是kA的属于$k\lambda$的特征向量；

(ii) $k+\lambda$是kI_n+A的特征值，且α是kI_n+A的属于$k+\lambda$的特征向量．

3．求证：Hermite阵的特征值全是实数．

4．求证：反对称阵的特征值是纯虚数或零．

5．求证：正交阵的特征值的模等于1．

6．求证：

(i) 对称正交阵的特征值是1或-1；

(ii) 反对称正交阵的特征值是$\sqrt{-1}$或$-\sqrt{-1}$．

7．设A、B都是n阶对称阵（或反对称阵），求证：$AB-BA$的特征值是纯虚数或零．

8．求上（下）三角阵的特征值．

9．设n阶阵A有s个不同的特征值$\lambda_1,\lambda_2,\cdots,\lambda_s$，$A$的属于$\lambda_i$的极大线性无关特征向量组中向量的个数为$n_i(i=1,2,\cdots,s)$．求证：若$A$有完全特征向量系，则

$$n_1+n_2+\cdots+n_s=n.$$

§2 方阵的相似、方阵相似于对角阵的条件

作为特征值的一个应用，本节主要讨论方阵相似于对角阵的条件．

一、方阵的相似

定义 设A、B同为数域K上的n阶方阵，称A与B（在数域K上）相似，如果存在K上的非异阵P，使得

$$B=P^{-1}AP.$$

A与B相似，记作$A\sim B$．

由定义不难证明下列命题

命题1 方阵的相似具有自反性($A\sim A$)、对称性(若$A\sim B$,

则 $B\sim A$)与传递性(若 $A\sim B$, $B\sim C$, 则 $A\sim C$)。

命题 2 相似的方阵有相同的秩、特征多项式、行列式。

【证明】 设 A 与 B 相似,即有非异阵 P 使得 $B=P^{-1}AP$,则由第三章定理 1.6 的推论可知 $r(A)=r(B)$。 又由 $B=P^{-1}AP$ 可得 $\lambda I-B=P^{-1}(\lambda I-A)P$,两边取行列式并应用乘法规则可得

$$|\lambda I-B|=|\lambda I-A|.$$

特别令 $\lambda=0$, 得到 $|B|=|A|$。 证毕。

命题 2 的逆命题未必成立。例如 $A=I_2$, $B=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$, 虽然 A 与 B 的秩、特征多项式、行列式都相同, 但 A 与 B 不相似。这是因为,对任何非异阵 P, $P^{-1}AP=I_2\neq B$。

二、方阵相似于对角阵的条件

在第三章 §4 例 2 中已经看到, 如能将方阵相似于一个对角阵,则对讨论方阵的性质将带来方便。但并非任何方阵都可以相似于对角阵,下面的定理将给出这一问题的一个充要条件。

定理 2.1 n 阶阵 A 可以相似于对角阵的充分必要条件是 A 有完全特征向量系。

【证明】 设 A 有完全特征向量系 $\alpha_1, \alpha_2, \cdots, \alpha_n$,

$$A\alpha_i=\lambda_i\alpha_i \quad (i=1, 2, \cdots, n),$$

令 $P=(\alpha_1, \alpha_2, \cdots, \alpha_n)$, 则

$$AP=(A\alpha_1, A\alpha_2, \cdots, A\alpha_n)=(\lambda_1\alpha_1, \lambda_2\alpha_2, \cdots, \lambda_n\alpha_n)$$

$$=(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}=P[\lambda_1, \lambda_2, \cdots, \lambda_n].$$

注意到 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关,故 P^{-1} 存在,上式即为

$$P^{-1}AP=[\lambda_1, \lambda_2, \cdots, \lambda_n].$$

这就证明了充分性。将上述证明过程倒推回去, 就是必要性的证明。其时,P 的各列是 A 的一个完全特征向量系,对角阵的主对角

元是与之相应的特征值。 证毕。

定理 2.1 的证明是构造性的。用于上节例 2 中的 A，可知非异阵

$$P=\begin{pmatrix}-1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 1 & 1\end{pmatrix}$$

可以使 A 相似于对角阵[1, 2, 2]。即

$$\begin{pmatrix}-1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 1 & 1\end{pmatrix}^{-1}\begin{pmatrix}1 & -1 & 1\\ 1 & 3 & -1\\ 1 & 1 & 1\end{pmatrix}\begin{pmatrix}-1 & 1 & 0\\ 1 & 0 & 1\\ 1 & 1 & 1\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}.$$

但上节例 3 的 A 不能相似于对角阵，因为它没有完全特征向量系。

推论1 设 n 阶阵 A 有 s 个不同的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$，$A$ 的属于 λ_i 的极大线性无关向量组中向量的个数为 n_i $(i=1, 2, \cdots, s)$，则 A 可以相似于对角阵的充分必要条件是

$$\sum_{i=1}^{s} n_i = n.$$

【证明】 利用上节定理 1.3 的推论可得。 证毕。

推论 2 如果 n 阶阵 A 有 n 个不同的特征值，则 A 必相似于对角阵。

【证明】 由假设 A 有 n 个不同的特征值，故由定理 1.2 可知，A 有完全特征向量系。再由定理 2.1 得到 A 必相似于对角阵。

证毕。

三、任意方阵相似于上三角阵

尽管方阵不一定都可以相似于对角阵，然而却可以相似于形状较简单的方阵，如三角阵。

定理 2.2 任一 n 阶复方阵 A 必可相似于一个上三角阵，且上三角阵的主对角元是 A 的 n 个特征值。

【证明】 仅需证明结论的前一半，后一半可利用命题 2 直

接由前一半推出。

对 n 作归纳。当 $n=1$ 时，定理显然正确。今设 α_1 是 n 阶复方阵 A 的属于特征值 λ_1 的特征向量

$$A\alpha_1=\lambda_1\alpha_1,$$

注意到 $\alpha_1\neq 0$，由第三章定理 4.1 可知，存在非异阵 P，以 α_1 为其第 1 列，即 $Pe_1=\alpha_1$ 或 $P^{-1}\alpha_1=e_1$。于是

$$P^{-1}APe_1=P^{-1}A\alpha_1=P^{-1}\lambda_1\alpha_1=\lambda_1P^{-1}\alpha_1=\lambda_1e_1.$$

上式表明，$P^{-1}AP$ 的第 1 列是 $(\lambda_1, 0, \cdots, 0)'$，故可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & \beta\\ 0 & A_1\end{pmatrix},$$

其中，β 是 $n-1$ 维行向量，A_1 是 $n-1$ 阶阵。由归纳法假设，存在 $n-1$ 阶非异阵 P_1，使得

$$P_1^{-1}A_1P_1=\begin{pmatrix}\lambda_2 & & *\\ & \lambda_3 & \\ & & \ddots \\ & & & \lambda_n\end{pmatrix},$$

记 $Q=\begin{pmatrix}1 & 0\\ 0 & P_1\end{pmatrix}$，$T=PQ$，则 T 是非异阵，且

$$T^{-1}AT=Q^{-1}(P^{-1}AP)Q=\begin{pmatrix}1 & 0\\ 0 & P_1^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & \beta\\ 0 & A_1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & P_1\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_1 & \delta\\ 0 & P_1^{-1}AP_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & & *\\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n\end{pmatrix}.$$

证毕。

定理 2.2 的证明方法是常用的。

推论 1 设 A 是 n 阶复方阵，λ_i 是 $|\lambda I_n-A|$ 的 k_i 重根 $(i=1, 2, \cdots, s)$，则

$$n-r(\lambda_iI_n-A)\leqslant k_i,\quad (i=1, 2, \cdots, s).$$

【证明】 为方便起见，不妨对 $i=1$ 证之。

由定理 2.2 及假设条件可知，存在非异阵 T，使得

$$T^{-1}AT=\begin{pmatrix}\lambda_1 & & & & & \\ & \lambda_1 & & & * & \\ & & \ddots & & & \\ & & & \lambda_1 & & \\ & & & & \lambda_2 & \\ & & & & & \ddots \\ & & & & & & \lambda_s\end{pmatrix}\begin{matrix}\}k_1\\ \\ \}n-k_1,\end{matrix}$$

其中,$\lambda_2,\cdots,\lambda_s$都不等于$\lambda_1$。于是

$$T^{-1}(\lambda_1 I_n-A)T=\begin{pmatrix}0 & & & & \\ & \ddots & & * & \\ & & 0 & & \\ & & & \lambda_1-\lambda_2 & \\ & & & & \ddots \\ & & & & & \lambda_1-\lambda_s\end{pmatrix}\begin{matrix}\}k_1\\ \\ \}n-k_1,\end{matrix}$$

由此可知

$$r(\lambda_1 I_n-A)=r(T^{-1}(\lambda_1 I_n-A)T)\geqslant n-k_1,$$

亦即

$$n-r(\lambda_1 I_n-A)\leqslant k_1.$$ 证毕。

推论 1 说明，属于λ_i的线性无关的特征向量的个数(称为几何重数)不会超过λ_i的重数(称为代数重数)。

推论 2 设$\lambda_1,\lambda_2,\cdots,\lambda_n$是$n$阶复方阵$A$的特征值，则对任何多项式$f(x)$，$f(A)$的$n$个特征值是$f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$。特别地,$A^k$的$n$个特征值是$\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k$($k$为任一非负整数)。

【证明】 根据定理 2.2，有

$$A\sim\begin{pmatrix}\lambda_1 & & * \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n\end{pmatrix},$$

进而易知,对任何多项式$f(x)$应有

$$f(A)\sim f\left(\begin{pmatrix}\lambda_1 & & * \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n\end{pmatrix}\right).$$

注意到

$$f\left(\begin{pmatrix}\lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}\right)$$

仍为一上三角阵，故其主对角元 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$ 就是它的特征值。又因相似的方阵有相同的特征多项式，故 $f(A)$ 的特征值也是 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$。 证毕。

习 题

1. 判断下列方阵能否相似于对角阵：

(i) $\begin{pmatrix}1+\sqrt{-1} & 1 & \sqrt{-1} \\ 0 & 1-\sqrt{-1} & 0 \\ 0 & 0 & \sqrt{-1}\end{pmatrix}$； (ii) $\begin{pmatrix}1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1\end{pmatrix}$；

(iii) $\begin{pmatrix}1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0\end{pmatrix}$； (iv) $\begin{pmatrix}-1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & -1\end{pmatrix}$。

2. 设

$$A=\begin{pmatrix}0 & -1 & 1 \\ 1 & 0 & 2 \\ -1 & -2 & 0\end{pmatrix},$$

(i) 求非异阵 P，使 $P^{-1}AP$ 是对角阵；

(ii) 求 A^{100}。

3. 求证：非零的幂零阵没有完全特征向量系。

4. 设 n 阶实方阵 A 满足 $A^2+A+I=O$。求证，A(在实数域上)没有完全特征向量系。

5. 设 λ_i 是 n 阶阵 A 的不同的特征值，且 λ_i 的重数为 $k_i(i=1, 2, \cdots, s)$。求证，A 可以相似于对角阵的充分必要条件是

$$n-\mathrm{r}(\lambda_i I_n - A)=k_i \qquad (i=1,2,\cdots,s)。$$

6. 设 n 阶阵 A 有 n 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，$n$ 阶阵 B 有 n 个不同的特征值 $\mu_1,\mu_2,\cdots,\mu_n$，且 $\lambda_1\lambda_2\cdots\lambda_n$ 是 $\mu_1, \mu_2, \cdots, \mu_n$ 的某一个排列。求证：存在 n 阶阵 P 与 Q，使 $A=PQ, B=QP$。

7. 设 A 是 n 阶非异阵。求证，A 的特征值 $\lambda_i \neq 0$； 且 A^{-1} 的特征值是

$$\frac{1}{\lambda_i} \quad (i=1,2,\cdots,n).$$

8. 求证

(i) 幂零阵的特征值只能是零;

(ii) 对合阵的特征值只能是1或-1;

(iii) 幂等阵的特征值只能是零或1.

9. 证明:若n阶实方阵A的n个特征值全是实数,则A必可(实)相似于一个实的上三角阵.

10. 设A是n阶阵,I_n-A的任一特征值λ_i的模$|\lambda_i|<1(i=1, 2, \cdots, n)$. 求证$0<|\det A|<2^n$,其中$|\det A|$表示$A$的行列式的模.

§3 方阵的特征多项式、特征多项式的降阶定理

一、特征多项式的展开式

给定n阶阵A后,怎样求A的特征多项式?我们有下面的基本展开式:

定理3.1 设A是数域K上的n阶阵,把A的特征多项式写成:

$$|\lambda I_n - A| = \lambda^n - a_1\lambda^{n-1} + a_2\lambda^{n-2} - \cdots + (-1)^{n-1}a_{n-1}\lambda + (-1)^n a_n, \qquad (1)$$

则a_i等于A的所有i阶主子式之和$(i=1, 2, \cdots, n)$.

【证明】 由第二章§6公式(3),

$$|\lambda I_n + B| = \lambda^n + b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + b_{n-1}\lambda + b_n \qquad (2)$$

其中b_i是B的所有i阶主子式之和$(i=1, 2, \cdots, n)$. 今取$B=-A$,则B的任一i阶主子式等于$(-1)^i$乘以A中相应位置上的i阶主子式,故$b_i=(-1)^i a_i(i=1, 2, \cdots, n)$. 把它代入(2)式即得(1)式. 证毕.

展开式(1)主要应用于理论推导,其中特别值得注意的是系数a_1与a_n.

定义 称方阵A的所有主对角元之和为A的迹,记为$\mathrm{tr}(A)$.

显然，

$$a_1 = \mathrm{tr}(A)。 \tag{3}$$

又若在(1)式中令 $\lambda=0$，则

$$a_n = |A|。 \tag{4}$$

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 n 阶阵 A 的特征值，则

$$|\lambda I - A| = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)。$$

由 Vieta 公式可知

$$a_1 = \lambda_1 + \lambda_2 + \cdots + \lambda_n,$$

$$a_n = \lambda_1\lambda_2\cdots\lambda_n,$$

于是可得下面的基本命题。

命题 1　设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 n 阶阵 A 的特征值，则

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i, \tag{5}$$

$$|A| = \prod_{i=1}^{n} \lambda_i。 \tag{6}$$

例 1　设 α 是实的 n 维向量，求 $\alpha\alpha'$ 的 n 个特征值。

【解】　因 $r(\alpha\alpha')=1$，故(1)式中的 $a_i=0$ $(i=2, 3, \cdots, n)$。所以

$$|\lambda I_n - \alpha\alpha'| = \lambda^n - a_1\lambda^{n-1} = \lambda^{n-1}(\lambda - a_1)。 \tag{7}$$

记 $\alpha = (c_1, c_2, \cdots, c_n)'$，则

$$\alpha\alpha' = \begin{pmatrix} c_1^2 & c_1c_2 & \cdots\cdots & c_1c_n \\ c_2c_1 & c_2^2 & \cdots\cdots & c_2c_n \\ \cdots & \cdots & \cdots\cdots & \cdots \\ c_nc_1 & c_nc_2 & \cdots\cdots & c_n^2 \end{pmatrix},$$

由(3)式，

$$a_1 = \mathrm{tr}(\alpha\alpha') = \sum_{i=1}^{n} c_i^2 = \alpha'\alpha,$$

所以(7)式化为

$$|\lambda I_n - \alpha\alpha'| = \lambda^{n-1}(\lambda - \alpha'\alpha)。$$

因此 $\alpha\alpha'$ 有 $n-1$ 个特征值是零，另外一个特征值是 $\alpha'\alpha$。

由方阵迹的定义,不难证明下列结论。

(i) $\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B)$;

(ii) $\mathrm{tr}(kA)=k\,\mathrm{tr}(A)$,($k$ 为任一数);

(iii) $\mathrm{tr}(AB)=\mathrm{tr}(BA)$;

(iv) $\mathrm{tr}(P^{-1}AP)=\mathrm{tr}(A)$。

例 2 设 A、C 是同阶方阵,且 $\mathrm{tr}(C)\neq 0$。求证矩阵方程 $AX-XA=C$ 无解。

【证明】 若不然,有方阵 B 使得 $AB-BA=C$,则

$$\mathrm{tr}(C)=\mathrm{tr}(AB-BA)=\mathrm{tr}(AB)-\mathrm{tr}(BA)=0,$$

与假设矛盾,故 $AX-XA=C$ 无解。 证毕。

由例 2 可知,对任何方阵 A、B,恒有 $AB-BA\neq I$。

二、特征多项式的降价定理

定理 3.2 设 A 是 $m\times n$ 阵,B 是 $n\times m$ 阵,且 $m\geqslant n$,则

$$|\lambda I_m-AB|=\lambda^{m-n}|\lambda I_n-BA|。\tag{8}$$

称(8)式为特征多项式的降阶定理。

【证明】 设 A 的秩为 r,则由第三章定理 1.7 可知,存在 m 阶非异阵 P 及 n 阶非异阵 Q,使

$$PAQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}。\tag{9}$$

令

$$Q^{-1}BP^{-1}=\begin{matrix} r \\ \\ \end{matrix}\overset{r}{\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}},\tag{10}$$

以(10)式右乘(9)式得

$$PABP^{-1}=\begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix};$$

以(10)式左乘(9)式得

$$Q^{-1}BAQ=\begin{pmatrix} B_1 & O \\ B_3 & O \end{pmatrix}。$$

因相似方阵有相同的特征多项式，故

$$|\lambda I_m - AB| = \begin{vmatrix} \lambda I_r - B_1 & -B_2 \\ O & \lambda I_{m-r} \end{vmatrix} = \lambda^{m-r}|\lambda I_r - B_1|,$$

$$|\lambda I_n - BA| = \begin{vmatrix} \lambda I_r - B_1 & O \\ -B_3 & \lambda I_{n-r} \end{vmatrix} = \lambda^{n-r}|\lambda I_r - B_1|,$$

由上两式显然可得(8)式。 证毕。

推论 1 设 A 是 $m\times n$ 阵，B 是 $n\times m$ 阵，则 AB 与 BA 的非零特征值全相同。

推论 2 设 A、B 为同阶方阵，则 AB 与 BA 有相同的特征多项式。

例 3 设 A 是 n 阶非异阵，α、β 是 n 维非零列向量。求证：$|\lambda A-\alpha\beta'|$ 有一个根为 $\beta'A^{-1}\alpha$，其它根全是零(当 $A=I_n$，$\alpha=\beta$ 时即得例 1)。

【证明】 因为

$$|\lambda A-\alpha\beta'| = |A|\,|\lambda I_n-(A^{-1}\alpha)\beta'|,$$

故由(8)式可得

$$|\lambda A-\alpha\beta'| = |A|\cdot\lambda^{n-1}(\lambda-\beta'A^{-1}\alpha).$$

因假设 A 是非异阵，即 $|A|\neq 0$，故 $|\lambda A-\alpha\beta'|$ 有一个根为 $\beta'A^{-1}\alpha$，其它 $n-1$ 个根为零。 证毕。

例 4 求 n 阶对称阵

$$A=\begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 0 \end{pmatrix}$$

的 n 个特征值及 A 的行列式。

【解】 因为

$$|\lambda I_n - A| = \begin{vmatrix} (\lambda+1)-1 & -1 & -1 & \cdots & -1 \\ -1 & (\lambda+1)-1 & -1 & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & -1 & \cdots & (\lambda+1)-1 \end{vmatrix}$$

$$= \left| (\lambda+1)I_n - \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \right|$$

$$= \left| (\lambda+1)I_n - \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} (1, 1, \cdots, 1) \right|,$$

由特征多项式的降阶定理可得

$$|\lambda I_n - A| = (\lambda+1)^{n-1}\left[(\lambda+1) - (1, 1, \cdots, 1)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}\right]$$

$$= (\lambda+1)^{n-1}(\lambda+1-n)。$$

所以，A 有 $n-1$ 个特征值是 -1，另一个特征值是 $n-1$。由(6)式，$|A| = (-1)^{n-1}(n-1)$。

例3、例4表明，利用特征多项式的降阶定理可以较方便地求出一个方阵的特征多项式。这一方法用于奇异阵更为奏效。这是因为，若 C 是秩为 r 的 n 阶奇异阵($r<n$)，则 C 有满秩分解：$C=AB$。其中 A 为 $n\times r$ 列满秩阵，B 为 $r\times n$ 行满秩阵。利用特征多项式的降阶定理可得

$$|\lambda I_n - C| = |\lambda I_n - AB| = \lambda^{n-r}|\lambda I_r - BA|。$$

当 r 较小时，$|\lambda I_r - BA|$ 极易求出(见例3)。对某些非异阵，有时也可将它化为奇异阵来处理(见例4)。

最后指出，为了应用上的方便，常将特征多项式的降阶定理改

写成

$$|\lambda I_m + AB| = \lambda^{m-n}|\lambda I_n + BA|. \tag{11}$$

其中，A、B 分别是 $m\times n$ 与 $n\times m$ 阵，且 $m\geqslant n$（即在(8)式中将 A 换成 $-A$）。

习　　题

1．设 n 阶阵 A 的秩 $r<n$．求证：

(i) $|\lambda I_n - A|$ 至少有 $n-r$ 重零(特征)根；

(ii) $|\lambda I_n - \mathrm{adj}A|$ 至少有 $n-1$ 重零（特征)根，其中 $\mathrm{adj}A$ 表示 A 的伴随阵．

2．求证：

(i) 若 A 是幂等阵，则 $r(A)=\mathrm{tr}(A)$；

(ii) 方阵

$$\begin{pmatrix} 1 & & * \\ & 1 & \\ * & & \ddots \\ & & & 1 \end{pmatrix}$$

必不是幂等阵．其中，$*$ 处的元素不全为零．

3．求 n 阶实镜象阵 $I_n-2uu'(u'u=1)$ 的 n 个特征值及它的迹与行列式．

4．如果 n 阶对合阵 A 不是单位阵，求证 $\mathrm{tr}(A)\leqslant n-2$．

5．设 A 是反对称阵．求证 $\mathrm{tr}(A)=0$．

6．设 $H=A+\sqrt{-1}B$ 是复方阵，A 与 B 分别是 H 的实部(矩阵)与虚部(矩阵)．当 H 是 Hermite 阵时，求证：

(i) $\mathrm{tr}(H)=\mathrm{tr}(A)$；

(ii) $\mathrm{tr}(H^2)>\mathrm{tr}(A^2)$．

7．设 n 阶阵 A 有 n 个特征值，其中 $n-1$ 个为零．求证 $|A-\mathrm{tr}(A)I_n|=0$．

8．如果对任何方阵 B，恒有 $\mathrm{tr}(AB)=0$，求证 $A=O$．

9．设 A 与 B 是同阶方阵．求证 $AB+B$ 与 $BA+B$ 有相同的特征多项式．

10．设 $m\times n$ 阵 A 的秩为 r，B 与 C 都是 $n\times m$ 阵，且 $m\geqslant n$．求证 AB 与 CA 至少有 $n-r$ 个为零的特征根．

11．设 A、B、X 为同阶方阵，且 $AX=XB$，$r(X)=r$。求证在复数域上 A 与 B 至少有 r 个相同的特征值。

12．对下列 n 阶对称阵 A，求出其 n 个特征值，并进而求出 $|A|$。

(i)
$$A=\begin{pmatrix} a_1^2+b_1^2 & a_1a_2+b_1b_2 & \cdots & a_1a_n+b_1b_n \\ a_1a_2+b_1b_2 & a_2^2+b_2^2 & \cdots & a_2a_n+b_2b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_1a_n+b_1b_n & a_2a_n+b_2b_n & \cdots & a_n^2+b_n^2 \end{pmatrix};$$

(ii)
$$A=\begin{pmatrix} 1+a_1^2 & a_1a_2 & \cdots & a_1a_n \\ a_2a_1 & 1+a_2^2 & \cdots & a_2a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_na_1 & a_na_2 & \cdots & 1+a_n^2 \end{pmatrix};$$

(iii)
$$A=\begin{pmatrix} a_1^2 & a_1a_2+1 & \cdots & a_1a_n+1 \\ a_2a_1+1 & a_2^2 & \cdots & a_2a_n+1 \\ \cdots & \cdots & \cdots & \cdots \\ a_na_1+1 & a_na_2+1 & \cdots & a_n^2 \end{pmatrix}.$$

§4 矩阵多项式、Hamilton–Cayley定理

设 $D=[d_1, d_2, \cdots, d_n]$，易知其特征多项式为

$$f_D(\lambda)=|\lambda I_n-D|=(\lambda-d_1)(\lambda-d_2)\cdots(\lambda-d_n).$$

根据第四章 §7 中有关的定义，

$$f_D(D)=(D-d_1I_n)(D-d_2I_n)\cdots(D-d_nI_n).$$

注意到对角阵 $D-d_iI_n$ 的第 i 个主对角元为零 $(i=1, 2, \cdots, n)$，可得

$$f_D(D)=O.$$

上式表明，对于对角阵 D，其特征多项式 $f_D(\lambda)$ 在 D 处的(矩阵)值恰为零。

对 n 阶上三角阵

$$J=\begin{pmatrix} 1 & a & & \\ & 1 & \ddots & \\ & & \ddots & a \\ & & & 1 \end{pmatrix},$$

其特征多项式为

$$f_J(\lambda)=(\lambda-1)^n,$$

故

$$f_J(J)=(J-I_n)^n=\begin{pmatrix} 0 & a & & \\ & 0 & \ddots & \\ & & \ddots & a \\ & & & 0 \end{pmatrix}^n \quad O.$$

又如对 Frobenius 阵

$$F=\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{pmatrix},$$

其特征多项式为

$$f_F(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n,$$

亦有

$$f_F(F)=F^n+a_1F^{n-1}+\cdots+a_{n-1}F+a_nI_n=O$$

(见第一章选做题 10)。

那么，对一般的方阵 A，是否都有 $f_A(A)=O$ 呢？回答是肯定的。为了证明这一结论，先引进矩阵多项式的概念。

定义 设 λ 为一符号，且对数域 K 上任一 s 阶阵 A，形式积 $A\lambda^j=\lambda^jA$（j 为任一正整数）。若 $A_0, A_1, \cdots, A_n$ 是数域 K 上的 s 阶阵，$A_n\neq O$（n 为非负整数），则称形式表达式

$$M(\lambda)=A_n\lambda^n+A_{n-1}\lambda^{n-1}+\cdots+A_1\lambda+A_0$$

为（数域 K 上的）一个 n 次矩阵多项式。

规定 A_0 代表 $A_0\lambda^0$。

容易看出，当 $s=1$ 时，矩阵多项式就是第四章中讨论的数域上的一元多项式。

矩阵多项式的相等、加法、乘法的定义类似于数域上的多项式，这里不再一一赘述。

若将矩阵多项式 $M(\lambda)$ 中的 λ 取作 K 中的变量，则 $M(\lambda)$ 可以表

示为

$$M(\lambda)=\begin{pmatrix} f_{11}(\lambda) & f_{12}(\lambda) & \cdots & f_{1s}(\lambda) \\ f_{21}(\lambda) & f_{22}(\lambda) & \cdots & f_{2s}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ f_{s1}(\lambda) & f_{s2}(\lambda) & \cdots & f_{ss}(\lambda) \end{pmatrix},$$

其中，$f_{ij}(\lambda)$是K上的次数不高于n的多项式（$i, j=1, 2, \cdots, s$）。上式是以K上的多项式为元素的s阶阵，通常称为多项式矩阵或λ-阵。

例如，

$$A(\lambda)=\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}\lambda^4+\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\lambda^3+\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$=\begin{pmatrix} \lambda^4+1 & -\lambda^4+\lambda^3+1 \\ 2\lambda^4+\lambda^3 & 3\lambda^4+1 \end{pmatrix}.$$

定义 设

$$M(\lambda)=A_n\lambda^n+A_{n-1}\lambda^{n-1}+\cdots+A_1\lambda+A_0$$
$$=\lambda^nA_n+\lambda^{n-1}A_{n-1}+\cdots+\lambda A_1+A_0,$$

其中A_i是s阶阵（$i=0, 1, 2, \cdots, n$）。若A是某一s阶阵，则记

$$M_R(A)=A_nA^n+A_{n-1}A^{n-1}+\cdots+A_1A+A_0,$$
$$M_L(A)=A^nA_n+A^{n-1}A_{n-1}+\cdots+AA_1+A_0.$$

一般说来，$M_R(A)\neq M_L(A)$。当$s=1$时，显然

$$M_R(A)=M_L(A),$$

与第四章§7中的定义一致。

与数域上多项式的余数定理（第四章§6）类似，在矩阵多项式中也有余数定理。

定理 4.1 设$M(\lambda)=A_n\lambda^n+A_{n-1}\lambda^{n-1}+\cdots+A_1\lambda+A_0$是$n$次矩阵多项式，$A_i$是$s$阶阵（$i=0, 1, \cdots, n$）。又设$A$是任一$s$阶阵，则

(i) 存在唯一的矩阵多项式$R(\lambda)$与唯一的数字矩阵R_1，使成立

$$M(\lambda)=R(\lambda)(\lambda I_s-A)+R_1, \tag{1}$$

其中

$$R_1 = M_R(A); \tag{2}$$

(ii) 存在唯一的矩阵多项式 $L(\lambda)$ 与唯一的数字矩阵 L_1，使成立

$$M(\lambda) = (\lambda I_s - A)L(\lambda) + L_1, \tag{3}$$

其中

$$L_1 = M_L(A). \tag{4}$$

【证明】 今证(i)，仿此可证(ii)。因为

$$\begin{aligned}
M(\lambda) &= A_n\lambda^n + A_{n-1}\lambda^{n-1} + \cdots + A_1\lambda + A_0 \\
&= A_n\lambda^{n-1}(\lambda I_s - A) + (A_nA + A_{n-1})\lambda^{n-1} \\
&\quad + A_{n-2}\lambda^{n-2} + \cdots + A_1\lambda + A_0 \\
&= [A_n\lambda^{n-1} + (A_nA + A_{n-1})\lambda^{n-2}](\lambda I_s - A) \\
&\quad + (A_nA^2 + A_{n-1}A + A_{n-2})\lambda^{n-2} + A_{n-3}\lambda^{n-3} \\
&\quad + \cdots + A_1\lambda + A_0 = \cdots\cdots \\
&= [A_n\lambda^{n-1} + (A_nA + A_{n-1})\lambda^{n-2} \\
&\quad + (A_nA^2 + A_{n-1}A + A_{n-2})\lambda^{n-3} + \cdots \\
&\quad + (A_nA^{n-1} + A_{n-1}A^{n-2} + \cdots + A_1)](\lambda I_s - A) \\
&\quad + (A_nA^n + A_{n-1}A^{n-1} + \cdots + A_1A + A_0).
\end{aligned}$$

令

$$\begin{aligned}
R(\lambda) = A_n\lambda^{n-1} &+ (A_nA + A_{n-1})\lambda^{n-2} + \cdots \\
&+ (A_nA^{n-1} + A_nA^{n-2} + \cdots + A_1), \\
R_1 = A_nA^n &+ A_{n-1}A^{n-1} + \cdots + A_1A + A_0,
\end{aligned}$$

即可得(1)、(2)两式。

再证唯一性。设矩阵多项式 $T(\lambda)$ 及数字矩阵 T_1，使成立

$$M(\lambda) = T(\lambda)(\lambda I_s - A) + T_1,$$

于是

$$[T(\lambda) - R(\lambda)](\lambda I_s - A) = R_1 - T_1.$$

如果 $T(\lambda) - R(\lambda) \neq 0$，则上式左边关于 λ 的次数不小于 1，而右边关于 λ 的次数为零，这是不可能的。所以 $T(\lambda) = R(\lambda)$，进而亦有 $T_1 = R_1$。

证毕。

推论 设 A 是 s 阶阵，$f(\lambda)=a_n\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$ 是数域 K 上的 n 次多项式，$T(\lambda)$ 是 $n-1$ 次矩阵多项式。如果

$$f(\lambda)I_s=(\lambda I_s-A)T(\lambda), \tag{5}$$

则

$$f(A)=O.$$

【证明】 显然可将 $f(\lambda)I_s$ 写成矩阵多项式，即

$$f(\lambda)I_s=\begin{pmatrix} f(\lambda) & & & \\ & f(\lambda) & & \\ & & \ddots & \\ & & & f(\lambda) \end{pmatrix}$$

$$=(a_nI_s)\lambda^n+(a_{n-1}I_s)\lambda^{n-1}+\cdots+a_0I_s\equiv M(\lambda).$$

由定理 4.1，存在矩阵多项式 $L(\lambda)$，使得

$$f(\lambda)I_s=M(\lambda)=(\lambda I_s-A)L(\lambda)+M_L(A);$$

另一方面，假设条件(5)可改写成如下形式：

$$f(\lambda)I_s=(\lambda I_s-A)T(\lambda)+O.$$

利用定理 4.1 中的唯一性可得 $M_L(A)=O$。但

$$M_L(A)=f(A)I_s=f(A),$$

所以 $f(A)=O$。 证毕。

下面的定理将回答本节开头提出的问题。

定理 4.2 (Hamilton–Cayley 定理) 设

$$f(\lambda)=|\lambda I_n-A|=\lambda^n+b_1\lambda^{n-1}+\cdots+b_{n-1}\lambda+b_n$$

是 n 阶阵 A 的特征多项式，则 $f(A)=O$。

【证明】 设 $\mathrm{adj}(\lambda I_n-A)$ 是 λI_n-A 的伴随阵，则

$$(\lambda I_n-A)\mathrm{adj}(\lambda I_n-A)=|\lambda I_n-A|I_n=f(\lambda)I_n{}^{①},$$

由推论即可得 $f(A)=O$。 证毕。

今后，为叙述方便起见，当 $f(A)=O$ 时，就称 A 满足 $f(\lambda)$，或称 $f(\lambda)$ 是以 A 为根的多项式。

例 1 设 n 阶非零阵 A 的特征值全是零，求证 A 是幂零阵。

【证明】 由假设可知 A 的特征多项式是 λ^n。根据 Hamilton–

① 此处 λ 理解为数域 K 中任一暂时固定的数，故仍可用第二章中有关伴随阵的结论。

Cayley 定理即得 $A^n=O$。 证毕。

例 2 设 A 是 n 阶非异阵，则 $A^{-1}=g(A)$，其中 $g(\lambda)$ 是一个 $n-1$ 次多项式。

【证明】 设 A 的特征多项式为

$$|\lambda I_n - A| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n,$$

由 Hamilton-Cayley 定理，

$$A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI_n = O.$$

因假设 A 是非异阵，故 $a_n=(-1)^n|A|\neq 0$，于是上式可化为

$$-\frac{1}{a_n}(A^{n-1}+a_1A^{n-2}+\cdots+a_{n-1}I_n)A = I_n,$$

这表明

$$A^{-1} = -\frac{1}{a_n}(A^{n-1}+a_1A^{n-2}+\cdots+a_{n-1}I_n)\equiv g(A),$$

其中，

$$g(\lambda) = -\frac{1}{a_n}(\lambda^{n-1}+a_1\lambda^{n-2}+\cdots+a_{n-1})$$

是一个 $n-1$ 次多项式。 证毕。

例 3 设

$$A=\begin{pmatrix} 1 & 0 & -1 \\ 0 & \omega & \sqrt{-2} \\ 0 & 0 & \omega^2 \end{pmatrix}, \qquad \omega = \frac{-1+\sqrt{-3}}{2}.$$

计算 A^{100}。

【解】 易见 A 的特征多项式是

$$(\lambda-1)(\lambda-\omega)(\lambda-\omega^2) = \lambda^3-1,$$

故由 Hamilton-Cayley 定理，

$$A^3 = I_3.$$

于是

$$A^{100} = (A^3)^{33}A = A.$$

习　　题

1. 对下列方阵 A，求 A^{100}。

(i) $A=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$; (ii) $A=\begin{pmatrix}-\sqrt{2}&1&2\\0&0&\sqrt{-2}\\0&0&\sqrt{2}\end{pmatrix}$.

2. 设 n 阶阵 A 的 n 个特征值全相同且不等于零

(i) 求证：$nA-\mathrm{tr}(A)I_n$ 是奇异阵；

(ii) 求：A^{-1}。

3. 设

$$A=\begin{pmatrix}1&0&0\\1&0&1\\0&1&0\end{pmatrix},$$

求证：$A^n=A^{n-2}+A^2-I_3\ (n\geqslant 3)$。

4. 用 Hamilton-Cayley 定理求 A 的逆阵：

$$A=\begin{pmatrix}\omega&\sqrt{-1}&1\\0&1&-1\\0&0&\omega^2\end{pmatrix},\qquad \omega=\frac{-1+\sqrt{-3}}{2}.$$

5. 用 Hamilton-Cayley 定理证明：若 n 阶阵 A 的所有不同的特征值是 $\lambda_1,\lambda_2,\cdots,\lambda_s$，它们的重数分别是 $k_1,k_2,\cdots,k_s$，则

$$\prod_{i=1}^{s}(A-\lambda_iI_n)^{k_i}=O.$$

6. 设 n 阶非异阵 A 的特征多项式是

$$|\lambda I_n-A|=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n,$$

求证：A^{-1} 必满足下面的 n 次多项式：

$$\lambda^n+a_{n-1}a_n^{-1}\lambda^{n-1}+\cdots+a_1a_n^{-1}\lambda+a_n^{-1}.$$

7. 记 $f(\lambda)=|\lambda I_n+A|=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$，求证：

$f(A)=2(A^n+a_2A^{n-2}+\cdots+a_nI_n)$　（n 为偶数时）；

$f(A)=2(a_1A^{n-1}+a_3A^{n-3}+\cdots+a_{n-2}A^2+a_nI_n)$　（n 为奇数时）。

§5 最小多项式

Hamilton-Cayley 定理说明，任意 n 阶阵 A 的特征多项式（n 次）必定以 A 为"根"。现在要问，是否存在次数较 n 低的多项式也能以 A 为根？这是可能的。例如，对 n 阶阵 $A=\alpha\alpha'$（其中，α 为实

的 n 维列向量,$n\geqslant 3$),有 $A^2=(\alpha'\alpha)A$。由此可知,多项式 $g(\lambda)=\lambda(\lambda-\alpha'\alpha)$以 A 为根,而 $g(\lambda)$的次数为 2,显然小于 n。

定义 称首一多项式 $m(\lambda)$ 为方阵 A 的最小多项式,如果 $m(\lambda)$是以 A 为根的多项式中次数最低者。

命题 1 方阵的最小多项式存在且唯一。

【证明】 存在性由 Hamilton-Cayley 定理保证。下面证明唯一性。设 $m_1(\lambda)$、$m_2(\lambda)$ 同是方阵 A 的最小多项式,若 $m_1(\lambda)\neq m_2(\lambda)$,则 $\varphi(\lambda)\equiv m_1(\lambda)-m_2(\lambda)\neq 0$。由定义,$m_1(\lambda)$与 $m_2(\lambda)$ 是同次的首一多项式,故

$$\deg(\phi(\lambda))<\deg(m_1(\lambda)).$$

又 $\phi(A)=m_1(A)-m_2(A)=O$,这说明 $m_1(\lambda)$ 不是 A 的最小多项式,与假设矛盾。于是有 $m_1(\lambda)=m_2(\lambda)$。 证毕。

今后,将 A 的最小多项式记为 $m_A(\lambda)$。

命题 2 A 的最小多项式整除任一个以 A 为根的多项式。

【证明】 设 $g(\lambda)$ 为任一个以 A 为根的多项式。若 $m_A(\lambda)\nmid g(\lambda)$,则由带余除法可知,存在 $q(\lambda)$以及 $r(\lambda)\neq 0$,

$$\deg(r(\lambda))<\deg(m_A(\lambda)),$$

使得

$$g(\lambda)=q(\lambda)m_A(\lambda)+r(\lambda),$$

故

$$g(A)=q(A)m_A(A)+r(A).$$

由假设 $g(A)=O$, $m_A(A)=O$,所以 $r(A)=O$。这说明 $m_A(\lambda)$不是 A 的最小多项式,与假设矛盾。从而有 $m_A(\lambda)\mid g(\lambda)$。 证毕。

推论 任一方阵的最小多项式必定整除其特征多项式。

推论表明,方阵的最小多项式的次数不会超过其特征多项式。而最小多项式,象特征多项式那样,亦有"以 A 为根"的性质。这就使得一些用 Hamilton-Cayley 定理解决的矩阵问题(如上节中求方阵的高次幂、非异阵的逆阵表示等),在改用最小多项式解决时更为方便。

下面介绍的命题,将有助于求出方阵的最小多项式。

命题 3 相似的方阵有相同的最小多项式。

【证明】 设 $B=P^{-1}AP$。由 §2 命题 2，对 A 的最小多项式 $m_A(\lambda)$，应有

$$m_A(B)=P^{-1}m_A(A)P.$$

因 $m_A(A)=O$，故 $m_A(B)=O$。利用上一命题可得 $m_B(\lambda)\mid m_A(\lambda)$。同理可证 $m_A(\lambda)\mid m_B(\lambda)$。又 $m_A(\lambda)$ 与 $m_B(\lambda)$ 都是首一多项式，故

$$m_A(\lambda)=m_B(\lambda).$$ 证毕。

命题 4 设分块对角阵 $A=[A_1, A_2]$，则

$$m_A(\lambda)=[m_{A_1}(\lambda), m_{A_2}(\lambda)]$$①。

【证明】 因 $m_A(A)=O$，即有

$$\begin{pmatrix} m_A(A_1) & O \\ O & m_A(A_2) \end{pmatrix}=O,$$

故

$$m_A(A_i)=O \quad (i=1, 2).$$

由命题 2 可知

$$m_{A_i}(\lambda)\mid m_A(\lambda) \quad (i=1, 2). \tag{1}$$

对 $m_{A_1}(\lambda)$、 $m_{A_2}(\lambda)$ 的任一公倍式 $\psi(\lambda)$，由 $m_{A_i}(A_i)=O$ 可推出 $\psi(A_i)=O \quad (i=1, 2)$，进而有

$$\psi(A)=\begin{pmatrix} \psi(A_1) & O \\ O & \psi(A_2) \end{pmatrix}=O.$$

再由命题 2 可知

$$m_A(\lambda)\mid \psi(\lambda), \tag{2}$$

综合(1)、(2)两式即得 $m_A(\lambda)=[m_{A_1}(\lambda), m_{A_2}(\lambda)]$。 证毕。

应用归纳法，可将命题 4 推广到一般的分块对角阵上去。

命题 5 设分块对角阵 $A=[A_1, A_2, \cdots, A_s]$，$s\geqslant 2$。则

---

① **定义** 称 $m(\lambda)$ 为 $f(\lambda)$ 与 $g(\lambda)$ 的最小公倍式，如果 $m(\lambda)$ 是 $f(\lambda)$ 与 $g(\lambda)$ 的公倍式，且对 $f(\lambda)$ 与 $g(\lambda)$ 的任一公倍式 $\psi(\lambda)$，都有 $m(\lambda)\mid\psi(\lambda)$。

由定义可证，两个多项式的最小公倍式之间最多相差一个常数因子。$f(\lambda)$ 与 $g(\lambda)$ 的最小公倍式(规定首项系数为 1)记为 $[f(\lambda), g(\lambda)]$。

$$m_A(\lambda)=[m_{A_1}(\lambda), m_{A_2}(\lambda), \cdots, m_{A_s}(\lambda)]①.$$

命题 6 （数域K上的）方阵（在K上）的特征值必定是其最小多项式（在K上）的根。

【证明】 用反证法。若λ_0是方阵A的特征值，而不是其最小多项式的根，则

$$(m_A(\lambda), \lambda-\lambda_0)=1.$$

于是，存在多项式$u(\lambda)$、$v(\lambda)$，使得

$$u(\lambda)m_A(\lambda)+v(\lambda)(\lambda-\lambda_0)=1,$$

从而有

$$u(A)m_A(A)+v(A)(A-\lambda_0 I)=I.$$

注意到$m_A(A)=0$，上式即为

$$v(A)(A-\lambda_0 I)=I,$$

两边取行列式可得

$$|A-\lambda_0 I|\neq 0.$$

这与λ_0是A的特征值相矛盾，故λ_0必定是A的最小多项式的根。

证毕。

进一步揭示特征多项式与最小多项式之间关系的 Frobenius 定理将在第六章中介绍。

习　题

1. 求下列方阵的最小多项式

(i) $\begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$;　　(ii) $\begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$;

(iii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

2. 举例说明，最小多项式相同的两个同阶方阵未必相似。

---

① 多个多项式的最小公倍式的定义与两个时相仿。$f_1(\lambda), f_2(\lambda), \cdots, f_s(\lambda)$的最小公倍式（规定首项系数为 1）记为$[f_1(\lambda), f_2(\lambda), \cdots, f_s(\lambda)]$。可以证明：

$$[f_1(\lambda), f_2(\lambda), \cdots, f_s(\lambda)]=[[f_1(\lambda), f_2(\lambda), \cdots, f_{s-1}(\lambda)], f_s(\lambda)], s\geqslant 2.$$

3. 求证任一方阵与其转置阵有相同的最小多项式。

4. 证明：秩为 1 的 n 阶阵($n\geqslant 2$)的最小多项式是 $\lambda^2-\mathrm{tr}(A)\lambda$。

5. 分别求出幂等阵与对合阵的最小多项式。

6. 设 n 阶阵 A 有两个不同的特征值：$\omega=\dfrac{-1+\sqrt{-3}}{2}$ 及 ω^2，且 A 的最小多项式的次数是 2。求证 I_n+A 是非异阵，并求 $(I_n+A)^{-1}$。

7. 证明：n 阶阵 A 是非异阵 $\Longleftrightarrow$ A 的最小多项式的常数项非零。

8. 设 $m(\lambda)$ 是 A 的最小多项式，$g(\lambda)$ 是一多项式。求证，

(i) 若 $(m(\lambda),g(\lambda))=d(\lambda)$，则 $r(d(A))=r(g(A))$；

(ii) $g(A)$ 是非异阵 $\Longleftrightarrow$ $g(\lambda)$ 与 $m(\lambda)$ 互素。

选 做 题

1. 求 n 阶循环阵 C 的特征值以及属于这些特征值的特征向量：

$$C=\begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_2 & c_3 & c_4 & \cdots & c_1 \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{pmatrix}$$

(提示：将 C 表示为

$$C=\sum_{k=0}^{n-1}c_k\begin{pmatrix} O & I_{n-1} \\ 1 & O \end{pmatrix}^k,$$

并求出 $\begin{pmatrix} O & I_{n-1} \\ 1 & O \end{pmatrix}$ 的特征值及特征向量)。

2. 证明(用于计算特征多项式的)Leverrier-Frame 公式 设 A 为 n 阶阵，若

$$A_0=I_n,$$

$$a_i=\frac{\mathrm{tr}(A_{i-1}A)}{i},\quad A_i=a_iI_n-A_{i-1}A\ (i=1,2,\cdots,n),$$

则

$$|\lambda I_n-A|=\lambda^n-a_1\lambda^{n-1}+a_2\lambda^{n-2}-\cdots+(-1)^{n-1}a_{n-1}\lambda+(-1)^na_n.$$

(提示：应用 Vieta 公式及 Newton 公式——第四章选做题 7.)

3. 记号同上题。求证：若 A 非异，则

$$A^{-1}=\frac{1}{a_n}A_{n-1}.$$

4．设 n 阶实方阵 A 的特征值全是实数，且 A 的所有 1 阶主子式之和、所有 2 阶主子式之和全等于零。求证 A 是 n 次幂零阵。

（提示：先证明 A 的特征值全为零）。

5．设 n 阶阵 A 的主对角元全是 1，且其特征值全是非负数。求证 $|A| \leqslant 1$。

（提示：应用 Cauchy 不等式，即若 $a_i \geqslant 0, i=1,2,\cdots,n$，则

$$\left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}} \leqslant \frac{1}{n}\left(\sum_{i=1}^{n} a_i\right).$$

6．设 $M=\begin{pmatrix} A & B \\ O & C \end{pmatrix}$ 是实正规阵，A 是方阵。求证：

$$r(M)=r(A)+r(C).$$

（提示：先证明 $B=O$）。

7．证明下述求特征多项式的 Крылов 方法：设 n 阶阵 A 的特征多项式为 $f_A(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$，若存在 n 维列向量 α，使得 $(A^{n-1}\alpha, A^{n-2}\alpha, \cdots, A\alpha, \alpha)$ 是非异阵，则

$$(a_1, a_2, \cdots, a_n)' = -(A^{n-1}\alpha, A^{n-2}\alpha, \cdots, A\alpha, \alpha)^{-1}A^n\alpha.$$

8．设 3 阶正交阵 A 的行列式等于 -1。求证：

$$\mathrm{tr}(A^2)=2\mathrm{tr}(A)+(\mathrm{tr}(A))^2.$$

9．设 $A=(a_{ij})_{n\times n}$，记 $g(\lambda)=\frac{1}{2}\left(|\lambda I_n+A|-|\lambda I_n-A|\right)$。求证：若 $g(A)$ 是非异阵，则 n^2 阶阵

$$G=\begin{pmatrix} A+a_{11}I_n & a_{12}I_n & \cdots & a_{1n}I_n \\ a_{21}I_n & A+a_{22}I_n & \cdots & a_{2n}I_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}I_n & a_{n2}I_n & \cdots & A+a_{nn}I_n \end{pmatrix}$$

也是非异阵。

10．设 $A=(a_{ij})_{n\times n}$ 是复方阵，记

$$R_i=\sum_{\substack{j=1 \\ j\neq i}}^{n}|a_{ij}|,$$

称 $G_i: |z-a_{ii}|=R_i$ 为 Гершгорин 圆 $(i=1,2,\cdots,n)$。证明：A 的任一特征值必落在某个 Гершгорин 圆的内部或边界上。

11．设 $A=(a_{ij})_{n\times n}$ 是复方阵，记 $S_i=\sum_{j=1}^{n}|a_{ij}|$，$T_i=\sum_{j=1}^{n}|a_{ji}|$ $(i=1,2,\cdots,n)$。若 λ 为 A 的任一特征值，求证

$$|\lambda|\leqslant\min(\max_i S_i,\max_i T_i).$$

12. 设多项式 $f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$. 求证：$f(\lambda)$ 的任一个根 λ_0 必满足不等式

$$|\lambda_0|\leqslant\min(\max(\sum_{i=1}^n|a_i|,1),\ \max_i(|a_i|+1)).$$

(提示：将 $f(\lambda)$ 看作它的友阵的特征多项式)。

13. 设 $A=(a_{ij})_{n\times n}$ 是复方阵，数 $d_i\neq 0(i=1,2,\cdots,n)$. 记

$$\tilde{R}_i=\sum_{j=1}^n|d_id_j^{-1}a_{ij}|,\ \tilde{T}_i=\sum_{j=1}^n|d_id_j^{-1}a_{ji}|\ \ (i=1,2,\cdots,n).$$

若 λ 为 A 的任一特征值，求证，

$$|\lambda|\leqslant\min(\max_i\tilde{R}_i,\max_i\tilde{T}_i).$$

14. 应用第 10 题的结论证明：

(i) 严格对角占优阵的特征值全不等于零；

(ii) 严格对角占优阵必是非异阵。(此即 **Lévy-Hadamard** 定理，见第三章 § 2 例 3)。

15. 应用第 10 题的结论证明：

(i) 严格对角占优实方阵的任一特征值的实部必大于零；

(ii) 严格对角占优实方阵的行列式大于零(此即第三章 §2 习题 8)；

(iii) 严格对角占优对称阵的任何主子式大于零。

16. 设 A 是 n 阶实方阵，且对任何 n 维非零实(列)向量 α，恒有 $\alpha' A\alpha<0$. 求证：A 的特征值的实部全小于零。

17. 求 n 阶 Jacobi 阵：

$$A=\begin{pmatrix}\alpha&\beta&&&\\ \gamma&\alpha&\beta&&\\ &\gamma&\ddots&\ddots&\\ &&\ddots&\alpha&\beta\\ &&&\gamma&\alpha\end{pmatrix}$$

的全部特征值。此处 α,β,γ 是任意复数，且 $\beta\neq 0,\gamma\neq 0$. (提示：令

$$\lambda-\alpha=a+b,\ \beta\gamma=ab,$$

则

$$0=|\lambda I_n-A|=\frac{a^{n+1}-b^{n+1}}{a-b}$$

(易证 $a\neq b$)，于是由 $a^{n+1}-b^{n+1}=0$ 及 $ab=\beta\gamma$ 可解出 a 与 b，从而求得 λ.)

18. 求 n 阶 "Fibonacci 矩阵"，

$$F=\begin{pmatrix} 1 & 1 & & & & \\ -1 & 1 & 1 & & 0 & \\ & -1 & 1 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & \ddots & 1 & 1 \\ & & & & -1 & 1 \end{pmatrix}$$

的全部特征值，并证明F的任何阶主子式全大于零。

19. 求证：n 阶阵

$$\begin{pmatrix} 0 & -1 & & & & \\ 1 & 0 & -1 & & 0 & \\ & 1 & 0 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & \ddots & 0 & -1 \\ & & & & 1 & 0 \end{pmatrix}$$

必相似于一个对角阵，并求出这个对角阵的具体形状。

20. 设A与B都是n阶阵，$f(\lambda)$、$m_A(\lambda)$与$m_B(\lambda)$分别是A的特征多项式、最小多项式与B的最小多项式，如果$m_A(\lambda)$与$m_B(\lambda)$互素，求证：$f(B)$是非异阵。

21. 设$f_A(\lambda)$与$f_B(\lambda)$分别是n阶阵A与B的特征多项式，且$f_A(B)$是非异阵，求证：$(f_A(\lambda), f_B(\lambda))=1$。

22. 设A、B、C都是n阶阵，且A的特征值全不相同，B的特征值亦全不相同，又设$f_A(\lambda)$是A的特征多项式，且$f_A(B)$是非异阵，求证：$\begin{pmatrix} A & C \\ O & B \end{pmatrix}$有完全的特征向量系。

第六章　方阵的相似标准形

上一章讨论了方阵相似于对角阵的条件，本章将讨论数域上任意方阵的相似标准形，把方阵化为最简单的形状。本章要介绍的标准形是：有理标准形（Frobenius 标准形）、Jacobson 标准形、Jordan 标准形，並初步讨论了这三种标准形，特别是有理标准形与 Jordan 标准形的应用。

§1　方阵的相似与 λ-阵的相抵

本章中如无特别声明，均以 K 表数域。如果 K 上的方阵 A 与 B（在 K 上）相似，即存在非异阵 P，使 $P^{-1}AP=B$，则对任一数 $\lambda\in K$，必有

$$P^{-1}(\lambda I-A)P=\lambda I-B, \tag{1}$$

即 A 与 B 的特征矩阵满足关系式(1)。反之，成立下面的

命题 1　设 A 与 B 是 K 上的同阶方阵，如果存在 K 上的非异阵 P、Q，使得

$$Q(\lambda I-A)P=\lambda I-B, \tag{2}$$

则 A 与 B 必（在 K 上）相似。

【证明】　比较矩阵多项式（λ-阵）(2) 的两边关于 λ 前的系数矩阵，即得

$$QP=I,\ QAP=B,$$

于是 $Q=P^{-1}$，$P^{-1}AP=B$，故 A 与 B（在 K 上）相似。　　证毕。

由上面的说明及命题 1 可知，A 与 B 是否相似的问题可归结为：对 A 与 B 的特征矩阵 $\lambda I-A$ 与 $\lambda I-B$，是否存在（数字）非异阵 Q 与 P，使成立(2)式，由于 P、Q 是一种特殊的 λ-阵，所以也可将

(2)式看作 λ-阵之间的一种等式，为了讨论(2)式存在的条件，先对 λ-阵作进一步的讨论。

定义 称下列 $n\times n$ λ-阵：

$$T_{ij}(\varphi(\lambda))=\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & \varphi(\lambda) & \cdots & 1 & \\ & & & & & \ddots \\ & & & & & & 1\end{pmatrix}$$

为第三种初等 λ-阵，其中 $\varphi(\lambda)$是 K 上的(关于 λ 的)多项式。

显然，$T_{ij}(\varphi(\lambda))$是通常第三种初等(数字)阵 $T_{ij}(k)$的推广，$k\in K$，且成立

$$T_{ij}(\varphi(\lambda))T_{ij}(-\varphi(\lambda))=T_{ij}(-\varphi(\lambda))T_{ij}(\varphi(\lambda))=I,$$

所以记

$$T_{ij}(\varphi(\lambda))^{-1}=T_{ij}(-\varphi(\lambda)),$$

称 $T_{ij}(\varphi(\lambda))^{-1}$ 为 $T_{ij}(\varphi(\lambda))$的逆阵。

如果将 K 上的第一、二两种初等阵 P_{ij}，$D_i(k)$，$k\neq 0$ 看作 λ-阵，分别称为第一、二种初等 λ-阵，则统称这三种初等 λ-阵为初等 λ-阵。

定义 称 $A(\lambda)$为**非异 λ-阵**，如果 $A(\lambda)$是有限个初等 λ-阵的乘积。

由于每一个初等 λ-阵都有逆阵，故显然可得

命题 2 对非异 λ-阵 $A(\lambda)$，必可找到非异 λ-阵 $B(\lambda)$，使

$$A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=I,$$

称 $B(\lambda)$为 $A(\lambda)$的逆阵，以 $B(\lambda)=A(\lambda)^{-1}$ 记之。

易知非异 λ-阵的行列式是 K 中的一个非零数。

定义 设 $A(\lambda)$与 $B(\lambda)$都是 λ-阵，它们的元素都是 K 上的多项式，如果存在非异 λ-阵 $P(\lambda)$、$Q(\lambda)$，它们的元素都是 K 上的多项式，使得

$$P(\lambda)A(\lambda)Q(\lambda)=B(\lambda),$$

则称 $A(\lambda)$与 $B(\lambda)$(在 $K[\lambda]$上)**相抵**，或称 $A(\lambda)$ 与 $B(\lambda)$ 是相抵

的。

显然，λ-阵的相抵是通常数字矩阵相抵概念的推广（不过此处讨论的是方阵）。今证本章的一个关键性结论。

定理 1.1 方阵 A 与 B 相似的充要条件是，$\lambda I-A$ 与 $\lambda I-B$（在 $K[\lambda]$ 上）相抵。

【证明】 必要性由（1）式即可看出。今证充分性。由假设 $\lambda I-A$ 与 $\lambda I-B$ 相抵，故存在非异 λ-阵 $P(\lambda)$ 与 $Q(\lambda)$，使成立

$$P(\lambda)(\lambda I-A)Q(\lambda)=\lambda I-B, \tag{3}$$

也即

$$P(\lambda)(\lambda I-A)=(\lambda I-B)Q(\lambda)^{-1}, \tag{4}$$

由第五章定理 4.1 可得

$$P(\lambda)=(\lambda I-B)L(\lambda)+L_1,$$

把它代入(4)式，移项並合并同类项即得

$$L_1(\lambda I-A)=(\lambda I-B)[Q(\lambda)^{-1}-L(\lambda)(\lambda I-A)],$$

上式左边的 λ 的次数不大于 1，故右边方括号内的 λ 的次数必须小于 1。也就是，它是一个数字矩阵，记它为 M，即

$$Q(\lambda)^{-1}-L(\lambda)(\lambda I-A)=M, \tag{5}$$

于是

$$L_1(\lambda I-A)=(\lambda I-B)M. \tag{6}$$

今证 M 是非异阵，把(5)式改写成：

$$I=MQ(\lambda)+L(\lambda)(\lambda I-A)Q(\lambda),$$

由(3)式，上式可改写为

$$I=MQ(\lambda)+L(\lambda)P(\lambda)^{-1}(\lambda I-B), \tag{7}$$

再由第五章定理 4.1，

$$Q(\lambda)=R(\lambda)(\lambda I-B)+R_1, \tag{8}$$

将(8)式代入(7)式，並合并同类项，得到

$$I=MR_1+[MR(\lambda)+L(\lambda)P(\lambda)^{-1}](\lambda I-B), \tag{9}$$

由(9)式可知，它的右边方括号内的 λ-阵必须为零，故(9)式应是：$MR_1=I$。这说明 M 是非异阵，且 $M=R_1^{-1}$，把它代入(6)式，即得

$$L_1(\lambda I-A)R_1=(\lambda I-B), \tag{10}$$

故由命题1即知A与B相似。　　　　　　　　　　　　　　证毕。

定理1.1为我们论证A与B是否相似提供了一个新的途径，即只要论证$\lambda I-A$与$\lambda I-B$是否相抵就可以了。因此，如能将$\lambda I-A$与$\lambda I-B$是否相抵搞清楚，也就解决了A与B是否相似的问题，由于这一基本想法，以下将围绕特征矩阵的相抵问题展开讨论。

§2　λ-阵的初等变换、特征矩阵的法式

与数字矩阵一样，对一般的λ-阵$A(\lambda)$，我们可以作它的三种初等变换：第一、二两种初等变换与数字矩阵的第一、二两种初等变换完全相同，而第三种初等变换则是以非零多项式$\varphi(\lambda)$乘$A(\lambda)$的第i行（或列）加到$A(\lambda)$的第j行（或列）上去。统称这三种初等变换为λ-阵的初等变换。

容易证明，对λ-阵$A(\lambda)$作λ-阵的初等变换后得到的λ-阵$B(\lambda)$等于$A(\lambda)$按照八字规则左（右）乘相应的初等λ-阵。这与通常的数字矩阵的初等变换完全一致。

对任一n阶阵A，由于找$\lambda I-A$相抵的λ-阵时要涉及一般的$n\times n$λ-阵的相抵，故先讨论有关$A(\lambda)\neq 0$的一些相抵结果。首先，恒可假设$A(\lambda)$的左上角元素非零，且其次数不大于$A(\lambda)$的其他非零元素的次数。这因为，如果$a_{ij}(\lambda)$是$A(\lambda)$中具有上述性质的元素，则经过$A(\lambda)$的行之间、列之间的对调，可把$a_{ij}(\lambda)$调到与$A(\lambda)$相抵的新的λ-阵$B(\lambda)$的左上角，再对$B(\lambda)$讨论就可以了。故以后恒设$A(\lambda)$的左上角元素$a_{11}(\lambda)\neq 0$，且$a_{11}(\lambda)$的次数不大于$A(\lambda)$的其他非零元素的次数。今证：

引理　如果λ-阵$A(\lambda)\neq 0$中至少有一个非零元素不能被$a_{11}(\lambda)$所整除，则必可找到一个与$A(\lambda)$相抵的λ-阵$L(\lambda)$，使$L(\lambda)$的左上角元素$l_{11}(\lambda)\neq 0$，且$l_{11}(\lambda)$的次数低于$a_{11}(\lambda)$的次数，并且$l_{11}(\lambda)$可整除$L(\lambda)$的所有元素。

【证明】　下面的证明是构造性的。先设$n\times n$λ-阵$A(\lambda)$的第

一行上的某些元素、例如$a_{1j}(\lambda)$不能被$a_{11}(\lambda)$整除，则

$$a_{1j}(\lambda)=a_{11}(\lambda)\varphi(\lambda)+b_{11}(\lambda),$$

其中$b_{11}(\lambda)\neq 0$，且$\deg(b_{11}(\lambda))<\deg(a_{11}(\lambda))$，作初等变换，

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & \cdots & a_{1j}(\lambda)\cdots \\ \cdots & \cdots & \cdots \end{pmatrix} \quad (-\varphi(\lambda))\ \text{第1列加到第}j\text{列}$$

$$\to\begin{pmatrix} a_{11}(\lambda) & \cdots & b_{11}(\lambda)\cdots \\ \cdots & & \cdots \end{pmatrix} \quad \text{(交换第1列与第}j\text{列)}$$

$$\to\begin{pmatrix} b_{11}(\lambda) & \cdots & a_{11}(\lambda)\cdots \\ \cdots & \cdots & \cdots \end{pmatrix}$$

$$\equiv B(\lambda)=(b_{ij}(\lambda))_{n\times n}.$$

（上述初等变换的箭头所表示的意义与数字矩阵的初等变换的箭头相同。）

不妨设$b_{11}(\lambda)$的次数不大于$B(\lambda)$的所有其他非零元素的次数（不然的话，把$B(\lambda)$的非零元素中的次数最低者调到左上角），此时，如果$b_{11}(\lambda)$已能整除$B(\lambda)$的其他元素，则引理已得证，如果$b_{11}(\lambda)$还不能整除$B(\lambda)$的其他元素，则总可假设$b_{11}(\lambda)$能整除$B(\lambda)$的第一行与第一列上的所有元素，但不能整除不在第一行与第一列上的某些其他元素，这因为，如果$b_{11}(\lambda)$不能整除某些$b_{1j}(\lambda)$或$b_{i1}(\lambda)$，则仿上法，可得一个与$B(\lambda)$相抵（因而也与$A(\lambda)$相抵）的λ-阵$C(\lambda)$，它的左上角元素$c_{11}(\lambda)$的次数比$b_{11}(\lambda)$的次数低，如果这个$c_{11}(\lambda)$还不能整除$c(\lambda)$的第一行与第一列的各个元素，则再继续用上面的方法，对$c(\lambda)$的第一行或第一列，或两者交替进行初等变换，由于每进行一次初等变换，左上角的元素至少降低一次，而原来的$a_{11}(\lambda)$的次数是有限的，故经过有限次初等变换后，总可找到与$A(\lambda)$相抵的λ-阵，使它的第一行与第一列的所有元素都能被左上角元素所整除，故不妨设$B(\lambda)$就是具有这种性质的λ-阵，即

$$b_{11}(\lambda) \mid b_{1j}(\lambda),\ b_{11}(\lambda) \mid b_{i1}(\lambda),\ i,\ j=1,\ 2,\ \cdots,\ n$$

而 $b_{11}(\lambda) \nmid b_{kl}(\lambda)$，$k\neq 1$，$l\neq 1$。记 $b_{k1}(\lambda)=b_{11}(\lambda)g(\lambda)$，对 $B(\lambda)$ 作如下的初等变换：

$$B(\lambda)=\begin{matrix}(-g(\lambda))\\ \downarrow\end{matrix}\begin{pmatrix} b_{11}(\lambda) & \cdots & b_{1l}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ b_{k1}(\lambda) & \cdots & b_{kl}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

$$\xrightarrow{(1)}\begin{pmatrix} b_{11}(\lambda) & \cdots & b_{1l}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & b_{kl}(\lambda)-b_{1l}(\lambda)g(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

$$\to\begin{pmatrix} b_{11}(\lambda) & \cdots & (1-g(\lambda))b_{1l}(\lambda)+b_{kl}(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & b_{kl}(\lambda)-b_{1l}(\lambda)g(\lambda) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}\equiv M(\lambda).$$

因由假设，$b_{11}(\lambda) \mid b_{1l}(\lambda)$，$b_{11}(\lambda) \nmid b_{kl}(\lambda)$，故

$$b_{11}(\lambda) \nmid ((1-g(\lambda)b_{1l}(\lambda)+b_{kl}(\lambda)),$$

于是对 $M(\lambda)$ 作初等变换：

$$M(\lambda)\to\begin{pmatrix} s_{11}(\lambda) & \cdots \\ \cdots & \cdots \end{pmatrix}\equiv s(\lambda),$$

而 $s_{11}(\lambda)$ 的次数比 $b_{11}(\lambda)$ 的次数更低，且 $s_{11}(\lambda)\neq 0$，于是又可用上述方法于 $s(\lambda)$。由于 $b_{11}(\lambda)$ 的次数是有限的，故最后总可找到一个与 $A(\lambda)$ 相抵的 λ-阵 $L(\lambda)$，使它的左上角元素 $l_{11}(\lambda)$ 非零，$l_{11}(\lambda)$ 不仅能整除 $L(\lambda)$ 的第一行与第一列的元素，且能整除 $L(\lambda)$ 的所有其他元素。　　证毕。

定理 2.1　K 上的 n 阶阵 A 的特征矩阵 $\lambda I_n - A$ 必可相抵于一个对角形 λ-阵。具体说，即可找到非异 λ-阵 $P(\lambda)$ 与非异 λ-阵 $Q(\lambda)$，使

$$\begin{aligned} &P(\lambda)(\lambda I_n - A)Q(\lambda) \\ &=[1,\ \cdots,\ 1,\ d_{k+1}(\lambda),\ d_{k+2}(\lambda),\ \cdots,\ d_n(\lambda)], \end{aligned} \tag{1}$$

此处 $d_i(\lambda)$都是非常数的首一多项式.且

$$d_i(\lambda)\mid d_{i+1}(\lambda) \qquad (i=k+1,k+2,\cdots,n-1).$$

【证明】 分两种情形讨论:

(i) A 中至少有一个非主对角元 $a_{ij}\neq 0(i\neq j)$, 例如 $a_{12}\neq 0$, 则作下列初等变换:

$$\lambda I_n - A = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots \\ -a_{21} & \lambda - a_{22} & \cdots \\ \cdots\cdots & & \end{pmatrix}$$

$$\xrightarrow{(-a_{12}^{-1})\to} \begin{pmatrix} -a_{12} & \lambda - a_{11} & \cdots \\ \lambda - a_{22} & -a_{21} & \cdots \\ \cdots\cdots & & \end{pmatrix}$$

$$\xrightarrow{-(\lambda - a_{22})} \begin{pmatrix} 1 & a_{12}^{-1}(a_{11}-\lambda) & \cdots \\ \lambda - a_{22} & -a_{21} & \cdots \\ \cdots\cdots & & \end{pmatrix}$$

$$\xrightarrow{(a_{32})} \begin{pmatrix} 1 & a_{12}^{-1}(a_{11}-\lambda) & \cdots \\ 0 & a_{12}^{-1}\lambda^2 + b\lambda + c & \cdots \\ -a_{32} & \cdots\cdots & \\ \cdots\cdots & & \end{pmatrix}$$

$$\to\cdots\to \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A(\lambda) & \\ 0 & & & \end{pmatrix}, \tag{2}$$

其中 $A(\lambda)$是 $n-1$ 阶 λ-阵,对 $A(\lambda)$用引理,可找到有限个 λ-阵的初等变换,把 $\lambda I_n - A$ 化为

$$(\lambda I_n - A)\to\cdots\to\begin{pmatrix} 1 & 0 \\ 0 & A(\lambda) \end{pmatrix}\to\cdots\to\begin{pmatrix} 1 & 0 & 0 \\ 0 & d_2(\lambda) & * \\ 0 & * & * \end{pmatrix}$$

$$\equiv\begin{pmatrix} 1 & 0 \\ 0 & B(\lambda) \end{pmatrix},$$

其中 $B(\lambda)=\begin{pmatrix} d_2(\lambda) & * \\ * & * \end{pmatrix}$ 是与 $A(\lambda)$ 相抵的 $n-1$ 阶 λ-阵，且 $d_2(\lambda)$能整除 $B(\lambda)$的所有元素，$d_2(\lambda)$ 已用第二种初等变换化为首一多项式．于是再用初等变换把 $B(\lambda)$的第一行与第一列全消为零，即

$$(\lambda I_n-A)\to\cdots\to\begin{pmatrix} 1 & 0 \\ 0 & B(\lambda) \end{pmatrix}\to\begin{pmatrix} 1 & 0 & 0 \\ 0 & d_2(\lambda) & 0 \\ 0 & 0 & C(\lambda) \end{pmatrix},$$

其中$\begin{pmatrix} d_2(\lambda) & 0 \\ 0 & C(\lambda) \end{pmatrix}$仍是与 $B(\lambda)$ 相抵的 $n-1$ 阶 λ-阵，而 $C(\lambda)$的任一元素是 $d_2(\lambda)$乘某一多项式加到 $B(\lambda)$的某一(i, j) 元后得到的元素，故 $n-2$ 阶 λ-阵 $C(\lambda)$的元素仍然被 $d_2(\lambda)$所整除，于是用上面同样的方法．逐个进行有限次初等变换，最后化为：

$$(\lambda I_n-A)\to\begin{pmatrix} 1 & & \\ & d_2(\lambda) & \\ & & C(\lambda) \end{pmatrix}\to\begin{pmatrix} 1 & & & \\ & d_2(\lambda) & & \\ & & d_3(\lambda) & \\ & & & M(\lambda) \end{pmatrix}$$

$$\to\begin{pmatrix} 1 & & & & \\ & d_2(\lambda) & & & \\ & & d_3(\lambda) & & \\ & & & \ddots & \\ & & & & d_n(\lambda) \end{pmatrix},$$

且 $d_i(\lambda)\mid d_{i+1}(\lambda)$．在上述变换过程中的 $d_2(\lambda)$, $d_3(\lambda)$, $\cdots$, $d_n(\lambda)$，可能有前面的若干个是 1，例如 $B(\lambda)$中只要有一个元素是非零数（非零的零次多项式），则必 $d_2(\lambda)=1$，对 $C(\lambda)$等也是如此，故一般可设 $d_i(\lambda)\neq 1$, $i=k+1, k+2, \cdots, n$。

又上述变换过程中的每一个 λ-阵 $A(\lambda)$、$B(\lambda)$、$C(\lambda)\cdots$都是非零阵，否则，例如 $A(\lambda)=0$，则由（2）式，存在非异 λ-阵 $P(\lambda)$、$Q(\lambda)$，使

$$P(\lambda)(\lambda I_n - A)Q(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

于是 $|P(\lambda)|\cdot|\lambda I_n - A|\cdot|Q(\lambda)| = 0$，这是不可能的。因为，$P(\lambda)$ 与 $Q(\lambda)$ 都是有限个初等 λ-阵的乘积，故 $|P(\lambda)| = p \neq 0$，$|Q(\lambda)| = q \neq 0$，又 $|\lambda I_n - A|$ 显然不等于零。故所求出的 $d_i(\lambda)$ 全不等于零。

(ii) 如果 A 的所有非主对角元全是零(即 A 是对角阵)，此时

$$(\lambda I_n - A) = [\lambda - a_{11}, \lambda - a_{22}, \cdots, \lambda - a_{nn}],$$

由引理，可将 $\lambda I_n - A$ 化为：

$$(\lambda I_n - A) \to \cdots \to \begin{pmatrix} 1 & * \\ * & A(\lambda) \end{pmatrix},$$

于是又化为(i)的情形。　　　　证毕。

定义　称(1)式右边的对角形 λ-阵

$$[1, 1, \cdots, 1, d_{k+1}(\lambda), d_{k+2}(\lambda), \cdots, d_n(\lambda)]$$

为 $\lambda I_n - A$(在 $K[\lambda]$ 上)的**法式**。其中

$$d_i(\lambda) | d_{i+1}(\lambda),\ i = k+1, k+2, \cdots, n-1.$$

例 1　求 $\lambda I_3 - A$(在 $Q[\lambda]$ 上)的法式，其中

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 3 & -2 & 0 \\ -1 & 1 & -1 \end{pmatrix}$$

(Q 是有理数域)。

【解】

$$\lambda I_3 - A = \begin{pmatrix} \lambda & -1 & 1 \\ -3 & \lambda+2 & 0 \\ 1 & -1 & \lambda+1 \end{pmatrix}$$

$$\xrightarrow{(-\lambda)(3)} \begin{pmatrix} 1 & -1 & \lambda+1 \\ -3 & \lambda+2 & 0 \\ \lambda & -1 & 1 \end{pmatrix}$$

$$
\begin{array}{c} -(\lambda+1) \\ (1) \end{array}
\to\begin{pmatrix} 1 & -1 & \lambda+1 \\ 0 & \lambda-1 & 3\lambda+3 \\ 0 & \lambda-1 & -\lambda^2-\lambda+1 \end{pmatrix}
$$

$$
(-3) \to\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 3\lambda+3 \\ 0 & \lambda-1 & -\lambda^2-\lambda+1 \end{pmatrix}
$$

$$
\to\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 6 \\ 0 & \lambda-1 & -\lambda^2-4\lambda+4 \end{pmatrix}
$$

$$
(6) \to\begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & \lambda-1 \\ 0 & -\lambda^2-4\lambda+4 & \lambda-1 \end{pmatrix}
$$

$$
-(\lambda-1) \to\begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 6(\lambda-1) \\ 0 & -\lambda^2-4\lambda+4 & 6(\lambda-1) \end{pmatrix}
$$

$$
\to\left(\begin{array}{c} 1 \\ 6 \end{array}\right)\to\begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & -\lambda^2-4\lambda+4 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix}
$$

$$
(\lambda^2+4\lambda-4)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\lambda^2-4\lambda+4 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix}
$$

$$
\to\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-1)(\lambda^2+4\lambda+2) \end{pmatrix}.
$$

习　题

1. 求下列方阵的特征矩阵的法式，其中

(i) $A=\begin{pmatrix}1&0\\0&2\end{pmatrix}$,　　(ii) $A=\begin{pmatrix}1&-1&-1\\0&-3&-1\\2&-2&0\end{pmatrix}$.

2. 设 λI_n-A 的法式是 $[1,\cdots,1,d_{k+1}(\lambda),d_{k+2}(\lambda),\cdots,d_n(\lambda)]$,

$d_j(\lambda)\neq 1,\ j=k+1,\ \cdots,\ n;\ d_i(\lambda)\mid d_{i+1}(\lambda),\ i=k+1,\ \cdots,\ n-1.$

求证：A 的特征多项式

$$f(\lambda)=|\lambda I_n-A|=d_{k+1}(\lambda)d_{k+2}(\lambda)\cdots d_n(\lambda).$$

3. 求证：数域 K 上的方阵 A 与它的转置阵 A'（在 K 上）相似。

4. 设 A 是行列式为 1 的 n 阶实对合阵，A_{ij} 是 $|A|$ 的元素 a_{ij} 的代数余子式，$i,\ j=1,\ 2,\ \cdots,\ n$，求证：A 与 B 相似，其中

$$B=\begin{pmatrix}A_{11}&A_{12}&\cdots&A_{1n}\\A_{21}&A_{22}&\cdots&A_{2n}\\\cdots&\cdots&\cdots&\cdots\\A_{n1}&A_{n2}&\cdots&A_{nn}\end{pmatrix}.$$

§3　不变因子、有理标准形及其应用

一、行列式因子与不变因子

为了进一步讨论特征矩阵的法式，引进一般的 λ-阵的行列式因子与不变因子的概念

定义　对某一 k，$1\leqslant k\leqslant n$，如果 n 阶 λ-阵 $A(\lambda)$ 的所有 k 阶子式的最高公因式（它是首一多项式）不等于零，则称它为 $A(\lambda)$ 的 k 阶**行列式因子**，并记为 $D_k(\lambda)$。

由行列式按照它的一列的展开公式显然可知，$D_{k-1}(\lambda)$ 整除 $D_k(\lambda)$，$2\leqslant k\leqslant n$，故可记

$$f_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)},\qquad 1\leqslant k\leqslant n,\tag{1}$$

其中约定 $D_0(\lambda)=1$。称 $f_k(\lambda)$ 为 $A(\lambda)$ 的第 k 个**不变因子**，$1\leqslant k\leqslant n$。

容易算出，n 阶阵 A 的特征矩阵 $\lambda I_n - A$ 的法式的行列式因子是，

$$D_1(\lambda) = D_2(\lambda) = \cdots = D_k(\lambda) = 1,$$

$$D_{k+1}(\lambda) = d_{k+1}(\lambda),\ D_{k+2}(\lambda) = d_{k+1}(\lambda)d_{k+2}(\lambda),\ \cdots,$$

$$D_n(\lambda) = d_{k+1}(\lambda)d_{k+2}(\lambda)\cdots d_n(\lambda).$$

于是由(1)式，$\lambda I_n - A$ 的法式有 n 个不变因子：

$$1,\ \cdots,\ 1,\ f_{k+1}(\lambda) = d_{k+1}(\lambda),\ f_{k+2}(\lambda) = d_{k+2}(\lambda),\ \cdots,$$

$$f_n(\lambda) = d_n(\lambda),$$

它们恰好是这个法式的 n 个主对角元。下面将证明，这些 $D_k(\lambda)$ 与 $d_k(\lambda)$ 也是 $\lambda I_n - A$ 的行列式因子与不变因子。为此先证：

命题 1 λ-阵 $A(\lambda)$ 经过 λ-阵的任何初等变换后，它的行列式因子不变，因而它的不变因子也不变，即相抵的 λ-阵有相同的行列式因子，因而也有相同的不变因子。

【证明】 就 $A(\lambda)$ 的行的初等变换证明即可，列变换可仿此证明。

1．对第二种初等变换，由于 $A(\lambda)$ 与 $D_i(c)A(\lambda)$，$c \neq 0$，的 k 阶子式或者相同，或者相差一个常数因子 c，但因行列式因子都是首一多项式，故 $A(\lambda)$ 与 $D_i(c)A(\lambda)$ 的行列式因子相同。

2．对第一种初等变换，由于 $A(\lambda)$ 与 $P_{ij}A(\lambda)$ 的 k 阶子式或者相同，或者只改变符号，自然不影响它们的行列式因子(因为它们都是首一多项式)。

3．对第三种初等变换，$A(\lambda)$ 与 $T_{ij}(\varphi(\lambda))A(\lambda)$ 的 k 阶子式有下列两种可能：在 $T_{ij}(\varphi(\lambda))A(\lambda)$ 中，凡同时包含它的第 i 行与第 j 行的那些 k 阶子式，以及那些不包含第 i 行的 k 阶子式自然都与 $A(\lambda)$ 的这些相应的 k 阶子式相同，而那些包含 $T_{ij}(\varphi(\lambda))A(\lambda)$ 的第 i 行但不包含它的第 j 行的 k 阶子式显然都可以拆成（对它们的第 i 行）$A(\lambda)$ 的一个 k 阶子式与 $\varphi(\lambda)$ 乘以 $A(\lambda)$ 的另一个 k 阶子式之和，故 $A(\lambda)$ 的 k 阶行列式因子 $D_k(\lambda)$ 应整除 $T_{ij}(\varphi(\lambda))A(\lambda)$ 的 k 阶行列式因子 $\Delta_k(\lambda)$，即 $D_k(\lambda) \mid \Delta_k(\lambda)$，但当 $A(\lambda)$ 经过第三种初等变换 $T_{ij}(\varphi(\lambda))$ 变成 λ-阵 $T_{ij}(\varphi\lambda))A(\lambda)$ 时，我们可继

续用 $T_{ij}(\varphi(\lambda))^{-1}=T_{ij}(-\varphi(\lambda))$把它变成 $A(\lambda)$，即

$$T_{ij}(-\varphi(\lambda))(T_{ij}(\varphi(\lambda))A(\lambda))=A(\lambda),$$

故由上面的讨论可知，$T_{ij}(\varphi(\lambda))A(\lambda)$的 k 阶行列式因子 $\Delta_k(\lambda)$应整除$T_{ij}(-\varphi(\lambda))(T_{ij}(\varphi(\lambda)))A(\lambda)=A(\lambda)$的 k 阶行列式因子 D_k，即$\Delta_k(\lambda)|D_k(\lambda)$，故 $\Delta_k(\lambda)$与 $D_k(\lambda)$两者最多差一个常数因子，但它们都是首一多项式，故 $D_k(\lambda)=\Delta(\lambda)$。　　证毕。

由命题 1 即知，λI_n-A 的法式的行列式因子(组)与不变因子(组)就是 λI_n-A 的行列式因子（组）与不变因子(组)，为了叙述简便起见，以后就称它们为A的行列式因子(组)与A的不变因子(组)。　　证毕。

命题 2　λI_n-A 的法式是(由 A)唯一确定的。

【证明】　如果

$$C(\lambda)=[1,\cdots,1,\mu_{s+1}(\lambda),\mu_{s+2}(\lambda),\cdots,\mu_n(\lambda)]$$

是 λI_n-A 的另一法式，$\mu_i(\lambda)|\mu_{i+1}(\lambda)$，$i=s+1,\cdots,n-1$，则由命题 1，$\lambda I_n-A$ 与 $C(\lambda)$ 应有相同的不变因子（组）。但 λI_n-A 的不变因子(组) 是 $1,\cdots,1,d_{k+1}(\lambda),d_{k+2}(\lambda),\cdots,d_n(\lambda)$，$d_i(\lambda)|d_{i+1}(\lambda)$，$i=k+1,k+2,\cdots,n-1$。 由于 $C(\lambda)$ 至少有 s 个不变因子是 1，因而当 $s>k$ 时，必有 $d_{k+1}(\lambda)=\cdots=d_s(\lambda)=1$，并且 $d_j(\lambda)=\mu_j(\lambda)$　(由命题 1)，$j=s+1,\cdots,n$。而当 $s<k$ 时，必有

$$\mu_{s+1}(\lambda)=\cdots=\mu_k(\lambda)=1,$$

并且 $\mu_i(\lambda)=d_i(\lambda)$，$i=k+1,\cdots,n$。故 λI_n-A 的任意两个法式相同，也即 λI_n-A 的法式是唯一的。　　证毕。

例 1　设 K 与 Ω 是两个数域，且 $K\subset\Omega$，又设 A 与 B 是 K 上的两个方阵，如果 A 与 B 在 Ω 上相似，则 A 与 B 必在 K 上相似。

【证明】　由假设，A 与 B 在 Ω 上相似，故由定理 1.1，λI_n-A 与 λI_n-B (在 $\Omega[\lambda]$上) 相抵。但因 A 与 B 的元素都取自 K 中，故化 λI_n-A 为法式的过程实际上不需涉及 $\Omega[\lambda]$中的元素，即可找到以 K 上的多项式为元素的非异 λ-阵 $P(\lambda)$与 $Q(\lambda)$，使

$$\begin{aligned}&P(\lambda)(\lambda I_n-A)Q(\lambda)\\&=[1,\cdots,1,d_{k+1}(\lambda),d_{k+2}(\lambda),\cdots,d_n(\lambda)],\end{aligned}\tag{2}$$

而$d_i(\lambda)\in K[\lambda]$, $i=k+1, k+2, \cdots, n$。同理，可找到元素取自$K[\lambda]$的非异λ-阵$R(\lambda)$与$T(\lambda)$，使

$$R(\lambda)(\lambda I_n-B)T(\lambda)$$
$$=[1, \cdots, 1, \mu_{s+1}(\lambda), \mu_{s+2}(\lambda), \cdots, \mu_n(\lambda)], \qquad (3)$$

而$\mu_i(\lambda)\in K[\lambda]$, $i=s+1, s+2, \cdots, n$。显然，(2)与(3)式可分别看作λI_n-A与λI_n-B在$\Omega[\lambda]$上的法式，由于λI_n-A与λI_n-B在$\Omega[\lambda]$上相抵，故由命题2, $k=s$（若$d_{k+1}(\lambda)\neq 1$），且$d_i(\lambda)=\mu_i(\lambda)$, $i=k+1, \cdots, n$。故由(2)与(3)即得

$$\lambda I_n-B=R(\lambda)^{-1}P(\lambda)(\lambda I_n-A)Q(\lambda)T(\lambda)^{-1}, \qquad (4)$$

这说明λI_n-A与λI_n-B在$K[\lambda]$上相抵，故由定理1.1, A与B在K上相似。　　证毕。

定理 3.1　n阶阵A与B在K上相似的充要条件是，它们有相同的不变因子(组)，或者它们有相同的行列式因子(组)。

【证明】　由定理1.1，只要证λI_n-A与λI_n-B在$K[\lambda]$上相抵的充要条件是，A与B有相同的不变因子(组)就可以了。

由命题1，必要性是显然的，又如果λI_n-A与λI_n-B有相同的不变因子(组)：$1, \cdots, 1, d_{k+1}(\lambda), d_{k+2}(\lambda), \cdots, d_n(\lambda)$，则由定理2.1，存在非异$\lambda$-阵$P(\lambda)$、$Q(\lambda)$、$R(\lambda)$、$T(\lambda)$，使

$$P(\lambda)(\lambda I_n-A)Q(\lambda)=[1, 1, \cdots, 1, d_{k+1}(\lambda), \cdots, d_n(\lambda)]$$
$$=R(\lambda)(\lambda I_n-B)T(\lambda),$$

即成立(4)式，故λI_n-A与λI_n-B相抵。充分性得证。证毕。

二、有理标准形(Frobenius 标准形)

定理3.1使我们有可能去建立与A相似的一个标准形。设

$$1, \cdots, 1, d_{k+1}(\lambda), d_{k+2}(\lambda), \cdots, d_n(\lambda)$$

是A的不变因子组，且$d_{k+i}(\lambda)$的次数$r_i\geqslant 1$, $d_{k+i}(\lambda)\in K[\lambda]$, $i=1, 2, \cdots, n-k$，则由$d_i(\lambda)|d_{i+1}(\lambda)$可知

$$r_{k+1}\leqslant r_{k+2}\leqslant\cdots\leqslant r_n, \quad \sum_{i=k+1}^{n} r_i=n.$$

设$F_1, F_2, \cdots, F_{n-k}$分别是$d_{k+1}(\lambda), d_{k+2}(\lambda), \cdots, d_n(\lambda)$的友阵，

则 F_i 分别是 r_{k+i} 阶阵，$i=1, 2, \cdots, n-k$。

定义 称分块对角阵：

$$F=[F_1, F_2, \cdots, F_{n-k}]$$

为 A 的**有理标准形**，或 **Frobenius 标准形**，而称 F_i 为 **Frobenius 块**，$i=1, 2, \cdots, n-k$。

定理 3.2 数域 K 上的任何 n 阶阵 A（在 K 上）必相似于它的有理标准形 F。

【证明】 由定理 3.1，只要证 A 与 F 有相同的不变因子（组）即可。任取 A 的不变因子 $d_{k+i}(\lambda)$，$1\leqslant i\leqslant n-k$，设

$$d_{k+i}(\lambda)=\lambda^{r_i}+a_{i1}\lambda^{r_i-1}+\cdots+a_{i,r_i-1}\lambda+a_{ir_i},$$

考虑 $d_{k+i}(\lambda)$ 的友阵 F_i：

$$F_i=\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_{ir_i} \\ 1 & 0 & \cdots & 0 & -a_{i,r_i-1} \\ 0 & 1 & \cdots & 0 & -a_{i,r_i-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & -a_{i1} \end{pmatrix},$$

易知 F_i 的第 r_i 个行列式因子是 $|\lambda I_{r_i}-F|=d_{k+i}(\lambda)$。而 F_i 的第 r_i-1 个行列式因子必是 1。这是因为 $\lambda I_{r_i}-F_i$ 的如下的 r_i-1 阶子式：

$$\begin{vmatrix} -1 & \lambda & & & \\ & -1 & \lambda & & * \\ & & -1 & & \\ & & & \ddots & \lambda \\ 0 & & & & -1 \end{vmatrix}=(-1)^{r_i-1}$$

又 F_i 的前 r_i-1 个行列式因子必也全是 1（因为 $D_i(\lambda)|D_{i+1}(\lambda)$），所以 $\lambda I_{r_i}-F_i$ 的不变因子（组）是，1，$\cdots$，1，$d_{k+i}(\lambda)$，于是由定理 2.1，存在 r_i 阶非异 λ-阵 $P_i(\lambda)$，$Q_i(\lambda)$，使

$$P_i(\lambda)(\lambda I_{r_i}-F)Q_i(\lambda)=[1, \cdots, 1, d_{k+i}(\lambda)], \quad 1\leqslant i\leqslant k,$$

故对所有 $i=1, 2, \cdots, n-k$，成立下式：

$$[P_1(\lambda), \cdots, P_{n-k}(\lambda)](\lambda I_n-A)[Q_1(\lambda), \cdots, Q_{n-k}(\lambda)]$$
$$=[1, \cdots, 1, d_{k+1}(\lambda), \cdots, 1, \cdots, 1, d_n(\lambda)],$$

显然可用第一种初等阵 P_{ij} 作行与列的对调，把上式右边的对角阵的 1 都调到前面 k 个位置上，于是 $\lambda I_n - F$ 与对角阵$[1, \cdots, 1, d_{k+1}(\lambda), \cdots, d_n(\lambda)]$相抵，所以 A 与 F 相似。 证毕。

例 2 如果 4 阶阵 A 的不变因子(组)是

$$1, \lambda+1, (\lambda+1)^3 = \lambda^3 + 3\lambda^2 + 3\lambda + 1,$$

则 A 的有理标准形是由两个 Frobenius 块构成的分块对角阵：

$$F = [F_1, F_2] = \left(\begin{array}{c|ccc} -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -3 \end{array}\right).$$

例 3 求

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 3 & -2 & 0 \\ -1 & 1 & -1 \end{pmatrix}$$

的有理标准形。

［解］ 在 §2 例 1 中已求出 A 的不变因子(组)是

$$1, 1, (\lambda-1)(\lambda^2+4\lambda+2) = \lambda^3 + 3\lambda^2 - 2\lambda - 2,$$

故 A 的有理标准形是一个 Frobenius 块所成的 3 阶阵：

$$F = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 2 \\ 0 & 1 & -3 \end{pmatrix},$$

于是 A 在有理数域上相似于 F。

方阵的有理标准形 F 的一个突出的优点是，它不仅在理论上是存在的，而且总可以具体求出它，正是由于这一点，所以无论在应用上以及理论推导中，它都有较大的作用，特别是一些用别的方法难以论证的结论，有时却可用它来解决(将在本节第三段及习题中初步看到这一点)。

三、有理标准形的应用

方阵的有理标准形有广泛的应用，今由它来证明著名的：

定理 3.3(Frobenius) 设数域K上的n阶阵A的特征多项式与最小多项式分别是$|\lambda I_n - A| = f(\lambda)$与$m(\lambda)$，则

(i) $m(\lambda) = d_n(\lambda)$；

(ii) $f(\lambda)$在K上的任一不可约因式都是$m(\lambda)$的因式。

此处$d_n(\lambda)$是A的最后一个不变因子。

【证明】 先证(i)。因为A在K上相似于它的有理标准形

$$F = [F_1, F_2, \cdots, F_{n-k}],$$

而相似矩阵有相同的最小多项式，故F的最小多项式也是$m(\lambda)$。另一方面，F的最小多项式是F_1、F_2、…、F_{n-k}的最小多项式的最小公倍式，而F_i的最小多项式是$d_{k+i}(\lambda)$，由于

$$d_{k+i}(\lambda) | d_{k+i+1}(\lambda),\ i = 1, \cdots, n-k-1,$$

所以F_1、F_2、…、F_{n-k}的最小多项式的最小公倍式就是$m(\lambda)$，于是$m(\lambda) = d_n(\lambda)$。

再证(ii)。设$f(\lambda)$在K上的标准分解式是：

$$f(\lambda) = p_1(\lambda)^{r_1} p_2(\lambda)^{r_2} \cdots p_s(\lambda)^{r_s}, \tag{5}$$

其中$p_i(\lambda)$对K皆不可约，$i = 1, 2, \cdots, s$。但

$$f(\lambda) = d_{k+1}(\lambda) d_{k+2}(\lambda) \cdots d_n(\lambda), \tag{6}$$

此处$d_{k+i}(\lambda)$是A的非常数不变因子，$i = 1, 2, \cdots, n-k$。由(5)式与(6)式可知，任一$p_i(\lambda)$至少要整除某一$d_{k+j}(\lambda)$，$1 \leqslant i \leqslant s$，$1 \leqslant j \leqslant n-k$，但$d_{k+j}(\lambda) | d_n(\lambda)$，$j = 1, \cdots, n-k$，故$p_i(\lambda) | d_n(\lambda)$。

证毕。

推论 1 A的特征多项式与最小多项式相等的充要条件是，$\lambda I_n - A$的第$n-1$个行列式因子$D_{n-1}(\lambda) = 1$。

这因为，由$m(\lambda) = d(\lambda) = \dfrac{f(\lambda)}{D_{n-1}(\lambda)}$即得推论 1。

推论 2 A的任一特征值必是A的最小多项式$m(\lambda)$的复根。

这因为，$f(\lambda)$在复数域上的不可约因式必为$\lambda - \lambda_i$，$i = 1, 2, \cdots, n$，而由定理 3.2 的(ii)，$(\lambda - \lambda_i) | m(\lambda)$，即$m(\lambda_i) = 0$，$i = 1, 2, \cdots, n$。

推论 3 如果 n 阶阵 A 的 n 个特征值全不相同，则 A 的特征多项式与最小多项式相等。

这由推论 2 可以明显地看出。

由于有理标准形应用的广泛性，故再举一综合运用之例以说明它的作用，将在本章的选做题中给出它更多的应用。

*例 4（Wonenburger-Djoković，1966～1967） 数域 K 上的 n 阶阵 A 可分解为 K 上两个对合阵乘积的充要条件是，A 为非异阵，且 A 与 A^{-1}（在 K 上）相似。

【证明】 必要性是明显的，因为若

$$A=B_1B_2,$$

而 B_1, B_2 都是 K 上的对合阵，即 $B_i^{-1}=B_i$, $i=1,2$，则上式可改写为

$$A=B_1(B_2B_1)B_1^{-1}=B_1A^{-1}B_1^{-1},$$

所以 A 与 A^{-1} 相似。

充分性，设 A 的不变因子(组)是 $1,\cdots,1,d_{k+1}(\lambda),d_{k+2}(\lambda),\cdots,d_n(\lambda)$，而 $d_i(\lambda)$ 都是非常数多项式，$i=k+1,k+2,\cdots,n$。设 $d_{k+1}(\lambda),d_{k+2}(\lambda),\cdots,d_n(\lambda)$ 的友阵分别是 $F_1,F_2,\cdots,F_{n-k}$，则由定理 3.2，存在非异阵 P，使

$$P^{-1}AP=[F_1,F_2,\cdots,F_{n-k}],$$

由假设 A 是非异阵，故

$$P^{-1}A^{-1}P=[F_1^{-1},F_2^{-1},\cdots,F_{n-k}^{-1}]=B.$$

由假设 A 与 A^{-1} 相似，故 A 与 B 相似，因而它们有相同的不变因子(组)。今对任一 F_i，$1\leqslant i\leqslant n-k$，为方便计，记这个暂时取定的 $F_i=F$。设 F 为 r 阶阵，把 F 如下分块：

$$F=\begin{pmatrix} 0 & a_r \\ I_{r-1} & \alpha \end{pmatrix},\ \alpha=(-a_{r-1},-a_{r-2},\cdots,-a_1)',$$

其中 α 是 $r-1$ 维列向量。由假设 A 是非异阵，故 $F_i=F$ 也是非异阵，于是 $a_r\neq 0$，且

$$F^{-1}=\begin{pmatrix} -\alpha a_r^{-1} & I_{r-1} \\ -a_r^{-1} & 0 \end{pmatrix},$$

而

$$\mu(\lambda)=|\lambda I_r-F^{-1}|=\begin{vmatrix} * & -I_{r-1} \\ a_r^{-1} & * \end{vmatrix}$$

中至少有一个 $r-1$ 阶子式是 $(-1)^{r-1}$，故 λI_r-F^{-1} 的第 $r-1$ 个行列式因子是 1，因而 F^{-1} 的不变因子(组)是 $1,\cdots,1,\mu(\lambda)$。

设 F_1^{-1}，F_2^{-1}，$\cdots$，F_{n-k}^{-1} 的特征多项式分别是 $\mu_{k+1}(\lambda)$，$\mu_{k+2}(\lambda)$，$\cdots$，$\mu_n(\lambda)$，则由刚才的讨论可知，F_i^{-1} 的不变因子(组)是 $1,\cdots,1,\mu_{k+i}(\lambda)$，$i=1,2,\cdots,n-k$，于是 λI_n-B 与 $[1,\cdots,1,\mu_{k+1}(\lambda),\mu_{k+2}(\lambda),\cdots,\mu_n(\lambda)]$ 相抵。今证 $1,\cdots,1,\mu_{k+1}(\lambda),\mu_{k+2}(\lambda),\cdots,\mu_n(\lambda)$ 就是 B 的不变因子(组)。设 $\lambda_{i1},\lambda_{i2},\cdots,\lambda_{ir_i}$ 是 F_i 的特征值，则 $\lambda_{i1}^{-1},\lambda_{i2}^{-1},\cdots,\lambda_{ir_i}^{-1}$ 是 F_i^{-1} 的特征值，$1\leqslant i\leqslant n-k$，于是

$$d_{k+i}(\lambda)=|\lambda I_{ri}-F_i|=(\lambda-\lambda_{i1})(\lambda-\lambda_{i2})\cdots(\lambda-\lambda_{iri}),\quad 1\leqslant i\leqslant n-k, \tag{7}$$

$$\mu_{k+i}(\lambda)=|\lambda I_{ri}-F_i^{-1}|=(\lambda-\lambda_{i1}^{-1})(\lambda-\lambda_{i2}^{-1})\cdots(\lambda-\lambda_{iri}^{-1}),\quad 1\leqslant i\leqslant n-k, \tag{8}$$

由于 $d_{k+i}(\lambda)|d_{k+i+1}(\lambda)$，故对任一 j，$1\leqslant j\leqslant r_i$，恒有

$$(\lambda-\lambda_{ij})|d_{k+i+1}(\lambda),$$

所以由(7)式与(8)式显然可得

$$(\lambda-\lambda_{ij}^{-1})|\mu_{k+i+1}(\lambda)\qquad 1\leqslant j\leqslant r_i,$$

即 $\mu_{k+i}(\lambda)|\mu_{k+i+1}(\lambda)$，$i=1,2,\cdots,n-k$。这说明 $1,\cdots,1,\mu_{k+1}(\lambda),\mu_{k+2}(\lambda),\cdots,\mu_n(\lambda)$ 确是 B 的不变因子(组)，所以它们也是 A 的不变因子组，故由法式的唯一性可得，

$$d_i(\lambda)=\mu_i(\lambda),\ i=k+1,k+2,\cdots,n.$$

另一方面，对取定的一个 F_i，设 $\deg F_i=r_i=r$，並记 F_i 与 F_i^{-1} 的友阵分别是

$$d_i(\lambda)=\lambda^r+a_1\lambda^{r-1}+a_2\lambda^{r-2}+\cdots+a_{r-1}\lambda+a_r, \tag{9}$$

$$\mu_i(\lambda)=\lambda^r+b_1\lambda^{r-1}+b_2\lambda^{r-2}+\cdots+b_{r-1}\lambda+b_r, \tag{10}$$

则由 $\mu_i(\lambda)=d_i(\lambda)$ 可知

$$a_j=b_j,\ j=1,2,\cdots,r, \tag{11}$$

但由(9)、(10)式以及 Vieta 公式，

$$b_r=(-1)^r\prod_{j=1}^{r}\lambda_j^{-1}=(-1)^r\left(\prod_{j=1}^{r}\lambda_j\right)^{-1}=a_r^{-1},$$

$$\begin{aligned}b_k&=(-1)^k\Sigma\lambda_{j1}^{-1}\lambda_{j2}^{-1}\cdots\lambda_{jk}^{-1}\\&=(-1)^k\left(\prod_{j=1}^{r}\lambda_j\right)^{-1}\Sigma\lambda_{s_1}\lambda_{s_2}\cdots\lambda_{s_{r-k}}\\&=a_r^{-1}a_{r-k},\quad k=1,2,\cdots,r-1,\end{aligned}$$

此处 $1\leqslant j_1<j_2<\cdots<j_k\leqslant r$；$1\leqslant s_1<s_2<\cdots<s_{r-k}\leqslant r$，故由(11)式可得：

$$a_r=a_r^{-1},\ a_k=a_r^{-1}a_{r-k},\ k=1,2,\cdots,r-1,\tag{12}$$

于是由(12)式得到

$$a_r=1,\ a_k-a_{r-k}=0,\ k=1,2,\cdots,r-1.\tag{13}$$

或者，

$$a_r=-1,\ a_k+a_{r-k}=0,\ k=1,2,\cdots,r-1.\tag{14}$$

由(13)式与(14)式可知

$$F_i=\begin{pmatrix}0 & \mp 1\\ I_{r-1} & \alpha\end{pmatrix},$$

並且当 $a_r=1$ 时有分解式：

$$F_i=\begin{pmatrix}0 & -1\\ I_{r-1} & \alpha\end{pmatrix}=\begin{pmatrix}-1 & 0\\ \alpha & J_{r-1}\end{pmatrix}\begin{pmatrix}0 & 1\\ J_{r-1} & 0\end{pmatrix},\tag{15}$$

其中

$$J=\begin{pmatrix} & & & 1\\ & & \cdot^{\cdot^{\cdot}} & \\ & 1 & & \\ 1 & & & \end{pmatrix},$$

易知

$$\begin{pmatrix}0 & 1\\ J_{r-1} & 0\end{pmatrix}=J_r$$

是对合阵，又由(13)式显然可知，$\alpha-J_{r-1}\alpha=0$，故

$$\begin{pmatrix}-1 & 0\\ \alpha & J_{r-1}\end{pmatrix}\begin{pmatrix}-1 & 0\\ \alpha & J_{r-1}\end{pmatrix}=\begin{pmatrix}1 & 0\\ -\alpha+J_{r-1}\alpha & I_{r-1}\end{pmatrix}=I_r,$$

即$\begin{pmatrix} -1 & 0 \\ \alpha & J_{r-1} \end{pmatrix}$也是对合阵。

同理，当 $a_r=-1$ 时，由(14)式可知 $\alpha+J_{r-1}\alpha=0$，由此可证

$$\begin{pmatrix} 1 & 0 \\ \alpha & J_{r-1} \end{pmatrix}^2=I_r,$$

且此时

$$F_i=\begin{pmatrix} 0 & 1 \\ I_{r-1} & \alpha \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ \alpha & J_{r-1} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ J_{r-1} & 0 \end{pmatrix}, \tag{16}$$

故由（15）式与（16）式可知，当 A 与 A^{-1} 相似时，A 的每一个 Frobenius 块 F_i 恒可分解为 K 上两个对合阵 $B_i^{(1)}$, $B_i^{(2)}$ 的乘积：

$$F_i=B_i^{(1)}B_i^{(2)},\ i=1,2,\cdots,n-k,$$

故由上式即得

$$\begin{aligned} A&=P[F_1,F_2,\cdots,F_{n-k}]P^{-1}\\ &=P[B_1^{(1)}B_1^{(2)},B_2^{(1)}B_2^{(2)},\cdots,B_{n-k}^{(1)}B_{n-k}^{(2)}]P^{-1}\\ &=(P[B_1^{(1)},B_2^{(1)},\cdots,B_{n-k}^{(1)}]P^{-1})\\ &\quad(P[B_1^{(2)},B_2^{(2)},\cdots,B_{n-k}^{(2)}]P^{-1})=B_1B_2, \end{aligned}$$

易知

$$B_1=P[B_1^{(1)},B_2^{(1)},\cdots,B_{n-k}^{(1)}]P^{-1},$$
$$B_2=P[B_1^{(2)},B_2^{(2)},\cdots,B_{n-k}^{(2)}]P^{-1},$$

都是 K 上的对合阵。 证毕。

不难看出，例 4 的充分性的证明是构造性的，用它可具体求出当 A 与 A^{-1} 相似时 A 的分解式。例如，有理数域上的 3 阶阵：

$$A=\begin{pmatrix} 0 & 3 & -3 \\ \frac{1}{3} & 0 & 1 \\ \frac{1}{3} & -\frac{7}{2} & \frac{7}{2} \end{pmatrix}$$

的不变因子是 1，1，$\lambda^3-\frac{7}{2}\lambda^2+\frac{7}{2}\lambda-1$，由此易知 A 与 A^{-1} 有相同的特征值，它们是 $1, a, a^{-1}(a\neq 1)$，故 A 与 A^{-1} 相似，且 A 的有

理标准形是

$$F=\begin{pmatrix}0 & 0 & 1\\ 1 & 0 & -\frac{7}{2}\\ 0 & 1 & \frac{7}{2}\end{pmatrix},$$

而变换阵 P 与其逆可取作

$$P=\begin{pmatrix}3 & 0 & 0\\ 0 & 1 & 1\\ 0 & 1 & 0\end{pmatrix},\qquad P^{-1}=\begin{pmatrix}3 & 0 & 0\\ 0 & 1 & 1\\ 0 & 1 & 0\end{pmatrix},$$

于是由例 5 的充分性证明过程((15)式)即得

$$A=\left(P\begin{pmatrix}1 & 0 & 0\\ -\frac{7}{2} & 0 & 1\\ \frac{7}{2} & 1 & 0\end{pmatrix}P^{-1}\right)\left(P\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}P^{-1}\right).$$

所以 A 分解为两个对合阵乘积的分解式是：

$$\begin{pmatrix}0 & 3 & -3\\ \frac{1}{3} & 0 & 1\\ \frac{1}{3} & -\frac{7}{2} & \frac{7}{2}\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -\frac{7}{6} & 1 & -1\end{pmatrix}\begin{pmatrix}0 & 3 & -3\\ \frac{1}{3} & 0 & 1\\ 0 & 0 & 1\end{pmatrix}.$$

习　　题

1. 用求行列式因子的方法求下列方阵 A 的特征矩阵的法式：

(i) $A=[1, 2, 3]$；　　(ii) $A=\begin{pmatrix}1 & 0 & 1\\ 2 & -1 & 1\\ 0 & 1 & -1\end{pmatrix}$.

2. 用 λ-阵的初等变换同时结合用行列式因子的方法，求下列方阵 A 的特征矩阵的法式：

(i) $A=\begin{pmatrix} 0 & 1 & -1 \\ 3 & -2 & 0 \\ -1 & 1 & -1 \end{pmatrix}$; (ii) $A=\begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & 1 \end{pmatrix}$;

(iii) $$A=\begin{pmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{pmatrix}.$$

3. 已知下列方阵的不变因子(组),写出A的有理标准形

(i) 不变因子(组)是$1, 1, 1, (\lambda-1), (\lambda-1)(\lambda+2), (\lambda-1)(\lambda+2)(\lambda+3)$;

(ii) 不变因子(组)是$1, 1, 1, \lambda+1, (\lambda+1)^2, (\lambda+1)^3$.

4. 求第1题中的各个A的有理标准形.

5. 设A是n阶幂零阵,求证:

(i) $\lambda I_n - A$的不变因子都是λ^l的形状,$0\leqslant l\leqslant n$;

(ii) A必相似于下面的分块对角阵:

$$[N_1, N_2, \cdots, N_s],$$

而N_i都是形如:

$$\begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$$

的n_i阶方阵.

6. 如果数域K上的n阶阵A的不变因子(组)是

$$1, \cdots, 1, f(\lambda)=|\lambda I_n - A|,$$

且$|A|=\pm1$(称这种A为K上的**么模阵**),求证:A必可分解为K上三个对合阵的乘积.

(提示:由题设A的有理标准形仅有一个Frobenius块F,设法把F写成三个对合阵的乘积.)

7. 如果数域K上的n阶么模阵A在K上有n个不同的特征值,则A必可分解为K上三个对合阵的乘积.

(提示:应用上一题的结论.)

§4 初等因子、Jacobson 标准形

一、初等因子

方阵A的有理标准形F虽然有很大优点，但当n较大时，F的每一个 Frobenius 块的阶数还可能很大，这是它的不足之处。本节将要把每一个 Frobenius 块分得更细，这样可以得到更精致的结果，为此引进A的初等因子的概念。

设$d_{k+1}(\lambda), d_{k+2}(\lambda), \cdots, d_n(\lambda)$是$A$的非常数不变因子，$d_i(\lambda) \in K[\lambda]$，$i=k+1, k+2, \cdots, n$。把$d_n(\lambda)$分解为$K$上的不可约因式的乘积：

$$d_n(\lambda)=p_1(\lambda)^{r_{n1}}P_2(\lambda)^{r_{n2}}\cdots P_t(\lambda)^{r_{nt}},$$

其中$p_i(\lambda)$在K上不可约，r_{ni}都是正整数，$i=1, 2, \cdots, t$。于是任一$d_j(\lambda)$可写成

$$d_j(\lambda)=p_1(\lambda)^{r_{j1}}p_2(\lambda)^{r_{j2}}\cdots p_t(\lambda)^{r_{jt}}, \quad k+1\leqslant j\leqslant n, \tag{1}$$

而r_{jl}是非负整数，且满足

$$0\leqslant r_{k+1,l}\leqslant r_{k+2,l}\leqslant\cdots\leqslant r_{nl}, \quad l=1, 2, \cdots, t$$

（因为$d_j(\lambda)|d_{j+1}(\lambda)$，$j=k+1, k+2, \cdots, n-1$）。

定义 当$r_{jl}\geqslant 1$时，称(1)中的每一个$p_l^{r_{jl}}$为A的（或λI_n-A的）**初等因子**。称A的所有初等因子的集合为A的**初等因子组**。

易知A的初等因子组是由A唯一确定的。

例1 设27阶阵A的不变因子组是：

$$\overbrace{1, \cdots, 1}^{24\text{个}1}, (\lambda-1)(\lambda^2-3),$$
$$(\lambda-1)^2(\lambda^2+\lambda+1)(\lambda^2-3)(\lambda^2-8)^2,$$
$$(\lambda-1)^2(\lambda^2+\lambda+1)^3(\lambda^2-3)(\lambda^2-8)^2,$$

则A在下列各自的基域上的初等因子组是：

1．在有理数域Q上共有10个初等因子，即

$(\lambda-1), (\lambda-1)^2, (\lambda-1)^2$; $\lambda^2-3, \lambda^2-3, \lambda^2-3$;

$(\lambda^2-8)^2, (\lambda^2-8)^2, \lambda^2+\lambda+1, (\lambda^2+\lambda+1)^3$.

2．在数域 $Q(\sqrt{2})$ 上共有 12 个初等因子，即

$$\lambda-1,\ (\lambda-1)^2,\ (\lambda-1)^2;\ \lambda^2-3,\ \lambda^2-3,\ \lambda^2-3;$$

$$(\lambda-2\sqrt{2})^2,\ (\lambda-2\sqrt{2})^2;$$

$$(\lambda+2\sqrt{2})^2,\ (\lambda+2\sqrt{2})^2,\ \lambda^2+\lambda+1,\ (\lambda^2+\lambda+1)^3.$$

3．在数域 $Q(\sqrt{3})$ 上共有 13 个初等因子，即

$$\lambda-1,\ (\lambda-1)^2,\ (\lambda-1)^2;\ \lambda-\sqrt{3},\ \lambda-\sqrt{3},\ \lambda-\sqrt{3};$$

$$\lambda+\sqrt{3},\ \lambda+\sqrt{3},\ \lambda+\sqrt{3};\ \ (\lambda^2-8)^2,\ (\lambda^2-8)^2;$$

$$\lambda^2+\lambda+1,\ (\lambda^2+\lambda+1)^3.$$

4．在实数域 R 上共有 15 个初等因子，即

$$\lambda-1,\ (\lambda-1)^2,\ (\lambda-1)^2;\ \lambda-\sqrt{3},\ \lambda-\sqrt{3},\ \lambda-\sqrt{3};$$

$$\lambda+\sqrt{3},\ \lambda+\sqrt{3},\ \lambda+\sqrt{3};\ (\lambda-2\sqrt{2})^2(\lambda-2\sqrt{2})^2;$$

$$(\lambda+2\sqrt{2})^2,\ (\lambda+2\sqrt{2})^2;\ \ \lambda^2+\lambda+1,\ (\lambda^2+\lambda+1)^3.$$

5．在复数域 C 上共有 17 个初等因子，即

$$\lambda-1,\ (\lambda-1)^2,\ (\lambda-1)^2;\ \ \lambda-\sqrt{3},\ \lambda-\sqrt{3},\ \lambda-\sqrt{3};$$

$$\lambda+\sqrt{3},\ \lambda+\sqrt{3},\ \lambda+\sqrt{3};\ (\lambda-2\sqrt{2})^2,\ (\lambda-2\sqrt{2})^2;$$

$$(\lambda+2\sqrt{2})^2,\ (\lambda+2\sqrt{2})^2;\ \ \lambda-\omega,\ (\lambda-\omega)^3;$$

$$\lambda-\omega^2,\ (\lambda-\omega^2)^3\left(\text{其中 } \omega=\frac{-1+\sqrt{-3}}{2}\right).$$

由上例可见，A 的初等因子组随着基域的不同而不同，基域愈扩大（因 $Q\subset Q\sqrt{2}\subset R\subset C$），初等因子愈多。

定理 4.1 数域 K 上的 n 阶阵 A 与 B 相似的充要条件是，它们有相同的初等因子组。

【证明】 必要性是明显的，因为 A 与 B 相似，故 $\lambda I-A$ 与 $\lambda I-B$ 有相同的不变因子(组)，因此它们有相同的初等因子组。

充分性。设 A 与 B 有相同的初等因子组，任取一个初等因子 $p_j(\lambda)^{r_{ij}}$，把所有 $p_j(\lambda)$ 的幂次方的那些初等因子按降幂排列：

$$p_j(\lambda)^{r_{nj}},\ p_j(\lambda)^{r_{n-1,j}},\ \cdots,\ p_j(\lambda)^{r_{k+1,j}};$$

$$r_{nj}\geqslant r_{n-1,j}\geqslant\cdots\geqslant r_{k+1,j}\geqslant 1,\quad j=1,2,\cdots,t, \tag{2}$$

作多项式

$$d_i(\lambda)=p_1(\lambda)^{r_{i1}}p_2(\lambda)^{r_{i2}}\cdots p_t(\lambda)^{r_{it}},\ i=k+1,\cdots,n-1,n, \tag{3}$$

易知 $d_i(\lambda)|d_{i+1}(\lambda)$, $i=k+1, \cdots, n-1$。所以由方阵的阶数 n 及它的初等因子组就唯一决定了 n 阶阵的不变因子(组)。今 A 与 B 都是 n 阶阵，且它们的初等因子组相同，所以 A 与 B 的不变因子(组)相同，因而 A 与 B 相似。 证毕。

定理 4.1 的充分性的证明过程实际上提供了一个由 A 的阶数以及它的初等因子组找出 A 的不变因子（组）的方法(通过 (2) 与 (3) 两步)。

二、Jacobson 块与 Jordan 块、Jacobson 标准形

设 $p(\lambda)^l=(\lambda^s+a_1\lambda^{s-1}+\cdots+a_s)^l$ 为数域 K 上的 n 阶阵 A(在 K 上)的任一初等因子，其中 $p(\lambda)=\lambda^s+a_1\lambda^{s-1}+\cdots+a_s$ 对 K 不可约，设 F 为 $p(\lambda)$ 的友阵，作 s 阶阵：

$$N=(e_s,\ 0,\ \cdots,\ 0)=\begin{pmatrix}0&0&\cdots&0\\ \vdots&\vdots&\cdots&\vdots\\ 0&0&\cdots&0\\ 1&0&\cdots&0\end{pmatrix},\tag{4}$$

再作 l 阶分块上三角阵

$$J=\begin{pmatrix}F'&N&&\\ &F'&N&\\ &&\ddots&N\\ &&&F'\end{pmatrix},\tag{5}$$

显然 J 是 ls 阶阵，而

$$F'=\begin{pmatrix}0&1&0&\cdots&0\\ 0&0&1&\cdots&0\\ \cdots&\cdots&\cdots&\cdots&\cdots\\ 0&0&0&\cdots&1\\ -a_s&-a_{s-1}&-a_{s-2}&\cdots&-a_1\end{pmatrix}\tag{6}$$

称(5)式的 J 为(相应于)初等因子 $p(\lambda)^l$ 的 Jacobson 块。

例 2 设有理数域 Q 上的方阵 A 的某一个初等因子为$(\lambda^3-\lambda^2$

$+1)^2$，则(相应于)$(\lambda^3-\lambda^2+1)^2$ 的 Jacobson 块为下面的 $2\times 3=6$ 阶阵：

$$J=\begin{pmatrix} F' & N \\ O & F' \end{pmatrix}=\begin{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \end{pmatrix}$$

$$=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}.$$

Jacobson 块的一个明显特点是，它的主对角线相邻的右上角对角线的元素全是 1，而其余上三角元素全是 0，于是它有一个 l_s-1 阶子式等于 1。

Jacobson 块的一个重要的特殊情形是，当不可约因式是一次式，此时 $s=1$，$N=(1)$，所以相应于$(\lambda-\lambda_i)^l$ 的 Jacobson 块是下面的 l 阶上三角阵：

$$J_l(\lambda_i)=\begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \lambda_i & 1 & \\ & & & \ddots & \ddots \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}$$

称 $J_l(\lambda_i)$为(相应于)初等因子$(\lambda-\lambda_i)^l$ 的 **Jordan 块**。

命题 1 $p(\lambda)^l$ 的 Jacobson 块 J 的特征矩阵 $\lambda I_{ls}-J$ 相抵于下面的 l_s 阶对角形 λ-阵：

$$[1, 1, \cdots, 1, p(\lambda)^l] \tag{7}$$

此处 $p(\lambda)=\lambda^s+a_1\lambda^{s-1}+\cdots+a_s$，且 $p(\lambda)$在 K 上不可约。

【证明】 因为 J 的第 l_s 个行列式因子 $D_{ls}(\lambda)$ 是：

$$D_{ls}(\lambda)=|\lambda I_{ls}-J|=\begin{vmatrix}\lambda I_s-F' & -N & & \\ & \lambda I_s-F' & -N & \\ & & \ddots & \ddots \\ & & & -N \\ & & & \lambda I_s-F'\end{vmatrix}$$

$$=|\lambda I_s-F'|^l=p(\lambda)^l,$$

其中 F 是 $p(\lambda)$ 的友阵，而 $\lambda I_{ls}-J$ 的右上角有一个 $ls-1$ 阶子式等于 $(-1)^{ls-1}$，故 $D_{ls-1}(\lambda)=1$，所以 J 的不变因子(组)是，$1,\cdots,1,p(\lambda)^l$，于是 $\lambda I_{ls}-J$ 与(7)式的对角形 λ-阵相抵。 证毕。

下面引进方阵 A 的 Jacobson 标准形。设 A 的不变因子（组）是 $1,\cdots,1,d_{k+1}(\lambda),d_{k+2}(\lambda),\cdots,d_n(\lambda)$，又设 $d_i(\lambda)$ 在数域 K 上的标准分解式是

$$d_i(\lambda)=p_1(\lambda)^{l_{i1}}p_2(\lambda)^{l_{i2}}\cdots p_t(\lambda)^{l_{it}},\ i=k+1,k+2,\cdots,n,$$

其中 $p_j(\lambda)$ 对 K 不可约。又设 $d_i(\lambda)$ 的分解式中那些不等于零的指数是 $l_{ij_1},l_{ij_2},\cdots l_{ij_s}$，而 $J_{ij_1},J_{ij_2},\cdots,J_{ij_s}$ 是(相应于)初等因子 $p_{j_1}(\lambda)^{l_{ij_1}},p_{j_2}(\lambda)^{l_{ij_2}},\cdots,p_{j_s}(\lambda)^{l_{ij_s}}$ 的 Jacobson 块，记

$$J_i=[J_{ij_1},J_{ij_2},\cdots,J_{ij_s}],\ i=k+1,k+2,\cdots,n \tag{8}$$

$$\tilde{J}=[J_{k+1},J_{k+2},\cdots,J_n], \tag{9}$$

称(9)式的 $\tilde{J}$ 为 A 的 **Jacobson 标准形**。

由 A 的 Jacobson 标准形的作法可知，$\tilde{J}$ 就是 A 的所有初等因子(所相应)的 Jacobson 块所成的分块对角阵，且可认为(8)式的 J_i 就是 A 的不变因子 $d_i(\lambda)$(所相应)的分块对角阵。

例 3 设实数域 R 上的 10 阶阵 A 的不变因子(组)是：

$$\overbrace{1,\cdots,1}^{8\text{ 个 }1},\ d_9(\lambda)=(\lambda-1)^2(\lambda^2+\lambda+1),$$
$$d_{10}(\lambda)=(\lambda-1)^2(\lambda^2+\lambda+1)^2,$$

求 A 的 Jacobson 标准形。

【解】 $d_9(\lambda)$ 的两个不可约因子(对 R)$(\lambda-1)^2$，$\lambda^2+\lambda+1$ (所相应)的 Jacobson 块分别是

$$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix},\ \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix},$$

故 $d_9(\lambda)$(所相应)的分块对角阵是

$$J_9=\begin{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} & \\ & \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix}\end{pmatrix},$$

而 $d_{10}(\lambda)$的不可约因子 $\lambda^2+\lambda+1$ 的平方(所相应)的 Jacobson 块是

$$\begin{pmatrix}\begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix} & \begin{pmatrix}0 & 0\\ -1 & 0\end{pmatrix}\\ & \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix}\end{pmatrix}$$

故 $d_{10}(\lambda)$(所相应)的 J_{10} 是

$$\boldsymbol{J}_{10}=\begin{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} & & \\ & \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix} & \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\\ & & \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix}\end{pmatrix},$$

所以 $\boldsymbol{A}$ 的 Jacobson 标准形是下面的 10 阶阵:

$$\tilde{J}=[J_9,\ J_{10}]$$

$$=\begin{bmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} & & & & \\ & \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix} & & & \\ & & \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} & & \\ & & & \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix} & \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\\ & & & & \begin{pmatrix}0 & 1\\ -1 & -1\end{pmatrix}\end{bmatrix}.$$

为了应用上的方便，通常把有相同初等因子(所相应)的各个 Jacobson 块紧接在一起，即

$$(\lambda-1)^2,\ (\lambda-1)^2;\ \lambda^2+\lambda+1,\ (\lambda^2+\lambda+1)^2,$$

因此相应的 Jacobson 块所成的分块阵可以写成

$$J=\begin{bmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}&&&&\\&\begin{pmatrix}1&1\\0&1\end{pmatrix}&&&\\&&\begin{pmatrix}0&1\\-1&-1\end{pmatrix}&&\\&&&\begin{pmatrix}0&1\\-1&-1\end{pmatrix}&\begin{pmatrix}0&0\\1&0\end{pmatrix}\\&&&&\begin{pmatrix}0&1\\-1&-1\end{pmatrix}\end{bmatrix},$$

显然 J 与 $\tilde{J}$ 是相似的，J 与 $\tilde{J}$ 仅是各个 Jacobson 块的排列次序不同而已。

一般地，设 n 阶阵 A 的初等因子组是

$$p_1(\lambda)^{l_{11}},\ p_1(\lambda)^{l_{21}},\ \cdots,\ p_1(\lambda)^{l_{s_1 1}};$$
$$p_2(\lambda)^{l_{12}},\ p_2(\lambda)^{l_{22}},\ \cdots,\ p_2(\lambda)^{l_{s_2 2}};\ \cdots\cdots;$$
$$p_t(\lambda)^{l_{1t}},\ p_t(\lambda)^{l_{2t}},\ \cdots,\ p_t(\lambda)^{l_{s_t t}}.$$

它们(相应)的 Jacobson 块依次记为

$$J_{11},\ J_{21},\ \cdots,\ J_{s_1,1};\ J_{12},\ J_{22},\ \cdots,\ J_{s_2,2};\ \cdots;$$
$$J_{1t},\ J_{2t},\ \cdots,\ J_{s_t,t}.$$

作分块对角阵：

$$J=[J_{11},\ J_{21},\ \cdots,\ J_{s_1,1},\ J_{21}J_{22},\ \cdots,\ J_{s_1,2},\ \cdots,\ J_{1t},\ J_{2t},\ \cdots,\ J_{s_t,t}],\tag{10}$$

也称 J 为 A 的 Jacobson 标准形。

因为一个方阵 A 的初等因子是由 A（及其阶数)唯一确定的，所以若不计 Jacobson 块的次序，则 A 的 Jacobson 标准形是(由 A)唯一确定的。

三、Jacobson 定理

本段将证明任何方阵A必相似于它的 Jacobson 标准形，为此先证

命题 2 设数域K上的n阶阵A的特征矩阵$\lambda I_n - A$相抵于对角形λ-阵：

$$B(\lambda) = [f_1(\lambda), f_2(\lambda), \cdots, f_n(\lambda)], \tag{11}$$

其中$f_n(\lambda)$与$f_i(\lambda)$都是K上的首一多项式，但$f_i(\lambda)$未必整除$f_{i+1}(\lambda)$, $i = 1, 2, \cdots, n-1$。如果$f_i(\lambda)$在K上的标准分解式是：

$$f_i(\lambda) = p_1(\lambda)^{l_{i1}} p_2(\lambda)^{l_{i2}} \cdots p_s(\lambda)^{l_{is}}, \quad i = 1, 2, \cdots, n,$$

则那些$l_{ij} \neq 0$所相应的$p_j(\lambda)^{l_{ij}}$的全体就是A的初等因子组。

【证明】 只要证明：(i) 那些$l_{ij} \neq 0$所相应的$p_j(\lambda)^{l_{ij}}$是A的初等因子；(ii) 所有这些$l_{ij} \cdot \deg p_j(\lambda)$的和等于$n$即可。

今先证(i)。不失一般性，可设

$$l_{11} \leqslant l_{21} \leqslant \cdots \leqslant l_{n1}, \tag{12}$$

否则可用λ-阵的初等变换调动$f_i(\lambda)$的次序，便可达到这一目的，今把(11)式的$B(\lambda)$写成：

$$B(\lambda) = [p_1(\lambda)^{l_{11}} g_1(\lambda), p_1(\lambda)^{l_{21}} g_2(\lambda), \cdots, p_1(\lambda)^{l_{n1}} g_n(\lambda)],$$

$$p_1(\lambda)^{l_{i1}} \nmid g_i(\lambda), \ 1 \leqslant i \leqslant n,$$

如果对某一k, $l_{k1} = 0$, $l_{k+1,1} \neq 0$，则由(12)式可知，

$$l_{11} = l_{21} = \cdots = l_{k1} = 0,$$

故此时$B(\lambda)$的第k阶行列式因子不包含$p_1(\lambda)$的幂次。今考察$B(\lambda)$的第$k+1$阶、$k+2$阶、…、n阶行列式因子，对任一r, $1 \leqslant r \leqslant n-1$，易知$B(\lambda)$的第$r$个行列式因子是

$$D_r(\lambda) = p_1(\lambda)^{l_{11}+l_{21}+\cdots+l_{r1}} \varphi(\lambda), \text{ 且 } p_1(\lambda) \nmid \varphi(\lambda),$$

而

$$D_{r+1}(\lambda) = p_1(\lambda)^{l_{11}+l_{21}+\cdots+l_{r1}+l_{r+1,1}} \psi(\lambda), \text{ 且 } p_1(\lambda) \nmid \psi(\lambda),$$

故$B(\lambda)$的不变因子是

$$d_{r+1}(\lambda) = \frac{D_{r+1}(\lambda)}{D_r(\lambda)} = p_1(\lambda)^{l_{r+1,1}} \nu(\lambda), \text{ 且 } p_1(\lambda) \nmid \nu(\lambda),$$

这说明，当 $l_{k+1,1}\neq 0$ 时，$B(\lambda)$的不变因子(也即A的不变因子)中含有 $p_1(\lambda)^{l_{k+1,1}}$, …, $p_1(\lambda)^{l_{n1}}$. 而且 $B(\lambda)$中含有 $p_1(\lambda)$的幂次的初等因子只能有这些。

与上面同样的论证方法可知，那些 $l_{ij}\neq 0$ 所相应的 $p_j(\lambda)^{l_{ij}}$, $2\leqslant j\leqslant n$, 都是A的某些不变因子的因子，且A的含 $p_j(\lambda)^{l_{ij}}$ 的初等因子只能有这些。

由 $f_i(\lambda)$的上述分解显然可知，这些 $l_{ij}\cdot\deg p_j(\lambda)$ 之和等于 n (由不变因子看出)，这证明了(ii)。 证毕。

由命题 2 可知，对角阵的初等因子(组)是：$\lambda-a_1$, $\lambda-a_2$, …, $\lambda-a_n$ (此处 a_1, a_2, …, a_n 是主对角元)。命题 2 还提供了一个求A的初等因子组的简便方法，即只要把 $\lambda I-A$ 相抵于某一个对角形 λ-阵，则它的主对角元的不可约因子的幂次方的全体就是A的初等因子组。我们常用此法先找出A的初等因子组，然后根据A的阶数 n，写出A的不变因子（组)。(从而也可求出 $\lambda I-A$ 的法式。)

定理 4.2(Jacobson) 数域K上的任何 n 阶阵A必(在K上)相似于它的 Jacobson 标准形 J。

【证明】 设A (在K上)的初等因子组是：

$$p_1(\lambda)^{l_{11}},\ p_1(\lambda)^{l_{21}},\ \cdots,\ p_1(\lambda)^{l_{s_1,1}};\ \cdots;$$
$$p_t(\lambda)^{l_{1t}},\ p_t(\lambda)^{l_{2t}},\ \cdots,\ p_t(\lambda)^{l_{s_t,t}},$$

它们(所相应)的 Jacobson 块是：

$$J_{11},\ J_{21},\ \cdots,\ J_{s_1,1};\ \cdots;\ J_{1t},\ J_{2t},\ \cdots,\ J_{s_t,t};$$

由命题 1，存在非异 λ-阵 $P_{ij}(\lambda)$, $Q_{ij}(\lambda)$ 使

$$\left.\begin{aligned}
&P_{11}(\lambda)(\lambda I_{11}-J_{11})Q_{11}(\lambda)=[1,\ \cdots\cdots,\ 1,\ p_1(\lambda)^{l_{11}}],\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&P_{s_1,1}(\lambda)(\lambda I_{s_1,1}-J_{s_1,1})Q_{s_1,1}(\lambda)=[1,\ \cdots\cdots,\ 1,\ p_1(\lambda)^{l_{s_1,1}}],\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&P_{1t}(\lambda)(\lambda I_{1t}-J_{1t})Q_{1t}(\lambda)=[1,\ \cdots\cdots,\ 1,\ p_t(\lambda)^{l_{1t}}],\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&P_{s_t,t}(\lambda)(\lambda I_{s_t,t}-J_{s_t,t})Q_{s_t,t}(\lambda)=[1,\ \cdots\cdots,\ 1,\ p_t(\lambda)^{l_{s_t,t}}],
\end{aligned}\right\}\quad(13)$$

其中 I_{ij} 是与 J_{ij} 同阶的单位阵。于是

$$
\begin{aligned}
& P(\lambda)(\lambda I_n - J)Q(\lambda) \\
& = P(\lambda)[\lambda I_{11} - J_{11}, \cdots, \lambda I_{s_1 1} - J_{s_1 1}, \cdots, \lambda I_{1t} \\
& \quad - J_{1t}, \cdots, \lambda I_{s_t t} - J_{s_t t}]Q(\lambda)
\end{aligned} \tag{14}
$$

此处记

$$P(\lambda) = [P_{11}(\lambda), \cdots, P_{s_1 1}(\lambda), \cdots, P_{1t}(\lambda), \cdots, P_{s_t t}(\lambda)],$$

$$Q(\lambda) = [Q_{11}(\lambda), \cdots, Q_{s_1 1}(\lambda), \cdots, Q_{1t}(\lambda), \cdots, Q_{s_t t}(\lambda)],$$

把(13)诸式代入(14)式，即得

$$
\begin{aligned}
& P(\lambda)(\lambda I_n - A)Q(\lambda) \\
& = [1, \cdots, 1, p_1(\lambda)^{l_{11}}, \cdots, 1, \cdots, 1, p_1(\lambda)^{l_{s_1 1}}, 1, \cdots, 1, \\
& \qquad p_t(\lambda)^{l_{1t}}, 1, \cdots, 1, p_t(\lambda)^{l_{s_t t}}],
\end{aligned}
$$

由命题 2 可知，A 与 J 的初等因子组相同，故由定理 4，A 与 J 相似。 证毕。

习　　题

1. 把下列方阵 A 的特征矩阵化成对角形 λ-阵，从而求出它们在各自基域 K 上的初等因子组：

(i) $A = \begin{pmatrix} 0 & 1 & -1 \\ 3 & -2 & 0 \\ -1 & 1 & -1 \end{pmatrix}$，$K = Q$；$Q(\sqrt{2})$；$Q(\sqrt{3})$．　（$Q$ 为有理数域）；

(ii) $A = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & 10 & 11 & -11 \end{pmatrix}$，$K = Q$。

2. 求 6 阶阵 $\left[1, -1, 2, 1, \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix}\right]$ 在 Q 上的 Jacobson 标准形。

3. 已知 11 阶阵 A 的不变因子（组）是：$\overbrace{1, \cdots\cdots, 1}^{9\text{个}1}$，$(\lambda+7)$ (λ^2-5)，$(\lambda+7)^2(\lambda^2-5)(\lambda^2-3)^2$，求 A 在下列各自基域上的 Jacobson 标准形：

(i) 有理数域 Q；(ii) $Q(\sqrt{3})$；(iii) $Q(\sqrt{5})$；(iv) 实数域 R。

4. 求证：数域K上的方阵A相似于对角阵的充要条件是，A的任何初等因子都是一次式。

5. 用 Jacobson 标准形证明：数域K上的方阵A相似于(K上)对角阵的充要条件是，A的最小多项式(在K上)有且只有单根。

6. 设有理数域上的 $2n$ 阶阵A的特征多项式的所有不可约因式是 x^2+x+1, x^2-2, 又设A的最小多项式是 4 次式，求证，A在复数域上必相似于对角阵。

§5 Jordan 标准形

本节讨论复数域C上的n阶阵A的 Jacobson标准形。因为复数域上的不可约因式必是一次式，故A的初等因子组都是$(\lambda-\lambda_j)^{l_{kj}}$的形状，因而每一 Jacobson 块都是 Jordan 块，于是由 Jacobson 定理显然可得：

定理 5.1 (Jordan) 任何复方阵A必相似于下面的分块对角阵：

$$J=[J(\lambda_1), J(\lambda_2), \cdots, J(\lambda_t)], \tag{1}$$

并且除了每一$J(\lambda_j)$的次序外，J是唯一的。此处$\lambda_1, \lambda_2, \cdots, \lambda_t$全不相同，而$J(\lambda_j)$是下列分块对角阵：

$$J(\lambda_j)=[J_{l_{1j}}(\lambda_j), J_{l_{2j}}(\lambda_j), \cdots, J_{l_{s_j j}}(\lambda_j)],$$
$$j=1, 2, \cdots, t, \tag{2}$$

其中的$J_{l_{kj}}(\lambda_j)$是下面的l_{kj}阶 Jordan 块：

$$J_{l_{kj}}(\lambda_j)=\begin{pmatrix}\lambda_j & 1 & & \\ & \lambda_j & 1 \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j\end{pmatrix},$$

$$k=1, 2, \cdots, s_j;\quad j=1, 2, \cdots; t, \tag{3}$$

称J为A的 **Jordan 标准形。**

例 1 求 4 阶阵：

$$A=\begin{pmatrix}-1 & 1 & 0 & 0\\ -1 & 0 & 1 & 0\\ 0 & 0 & -1 & 1\\ 0 & 0 & -1 & 0\end{pmatrix}$$

的 Jordan 标准形。

【解】 因为特征矩阵

$$\lambda I_4-A=\begin{pmatrix}\lambda+1 & -1 & 0 & 0\\ 1 & \lambda & -1 & 0\\ 0 & 0 & \lambda+1 & -1\\ 0 & 0 & 1 & \lambda\end{pmatrix}$$

的 4 阶行列式因子显然是：

$$D_4=|\lambda I_4-A|=(\lambda^2+\lambda+1)^2,$$

而 λI_4-A 又有一个 3 阶子式：

$$\begin{vmatrix}-1 & 0 & 0\\ \lambda & -1 & 0\\ 0 & \lambda+1 & -1\end{vmatrix}=(-1)^3=-1。$$

故 $D_3=1$，于是 A 的不变因子(组)是：$1, 1, 1, (\lambda^2+\lambda+1)^2$，所以 A 的初等因子组是：$(\lambda+\omega)^2$，$(\lambda+\omega^2)^2$，其中

$$\omega=\frac{-1+\sqrt{-3}}{2}。$$

由此可知 A 的 Jordan 标准形是：

$$J=[J_2(\omega), J_2(\omega^2)]=\begin{pmatrix}\omega & 1 & 0 & 0\\ 0 & \omega & 0 & 0\\ 0 & 0 & \omega^2 & 1\\ 0 & 0 & 0 & \omega^2\end{pmatrix}. \tag{4}$$

例1说明，有时未必要用 λ-阵的初等变换，而直接用求 λI_n-A 的行列式因子来找出 A 的初等因子乃是方便的。

Jordan 标准形在矩阵分解理论中也有它的应用，下面两例便是。

例 2 (Voss) 数域 K 上的任何方阵必可分解为复数域上的两个对称阵的乘积。

所谓复对称阵 S 指的是，满足 $S=S'$ 的复方阵，它未必是 Hermite 阵。例如$\begin{pmatrix} 1 & \sqrt{-1} \\ \sqrt{-1} & \sqrt{-1} \end{pmatrix}$是复对称阵，它显然不是 Hermite 阵。

【证明】 因为 A 可看作复数域上的方阵，故 A 可写成：

$$A=P[J(\lambda_1),J(\lambda_2),\cdots,J(\lambda_t)]P^{-1}=PJP^{-1}, \tag{5}$$

而分块对角阵 $J(\lambda_j)$ 中的每一个 Jordan 块 $J_{ij}(\lambda_j)$ 必可分解为下面两个复对称阵的乘积：

$$J_{ij}(\lambda_j)=\begin{pmatrix} \lambda_j & 1 & & \\ & \lambda_j & 1 \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{pmatrix}$$

$$=\begin{pmatrix} & & 1 & \lambda_j \\ & 1 & \lambda_j & \\ 1 & \lambda_j & & \\ \lambda_j & & & \end{pmatrix}\begin{pmatrix} & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \\ & 1 & & \\ 1 & & & \end{pmatrix}$$

$$=Q_{ij}^{(1)}Q_{ij}^{(2)},$$

故 J 可分解为下面两个复对称阵的乘积：

$$J=[Q_{11}^{(1)},\cdots,Q_{s_11}^{(1)},\cdots,Q_{1t}^{(1)},\cdots,Q_{s_tt}^{(1)}]\cdot[Q_{11}^{(2)},\cdots,$$
$$Q_{s_11}^{(2)},\cdots,Q_{1t}^{(2)},\cdots,Q_{s_tt}^{(2)}]=Q_1Q_2,$$

于是由(5)式，A 可分解为下面两个复对称阵的乘积：

$$A=PJP^{-1}=(PQ_1P')((P^{-1})'Q_2P^{-1}).$$ 证毕。

例 3 数域 K 上的任何方阵 A 必有分解式：

$$A=S+N, \tag{6}$$

并且 $SN=NS$。其中 N 是幂等阵，而 s (复)相似于分块对角阵：

$$[\lambda_1I_{n_1},\lambda_2I_{n_2},\cdots,\lambda_sI_{n_s}],$$

称 S 为半单纯阵。

【证明】 因为 $A=P[J(\lambda_1), J(\lambda_2), \cdots, J(\lambda_s)]P^{-1}$，而 $J(\lambda_i)$ 中的每一个 Jordan 块显然可分解为：

$$J_{ij}(\lambda_j)=\begin{pmatrix}\lambda_j & 1 & & \\ & \lambda_j & 1\ \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_j & & & \\ & \lambda_j & & \\ & & \ddots & \\ & & & \lambda_j\end{pmatrix}+\begin{pmatrix}0 & 1 & & & \\ & 0 & 1 & \ddots & \\ & & 0 & & 1 \\ & & & \ddots & 0\end{pmatrix}$$

$$=\lambda_j I_{ij}+N_{ij},$$

而 N_{ij} 是幂零阵。设 $J(\lambda_j)$为 n_j 阶阵，则 $J(\lambda_j)$有分解式：

$$J(\lambda_j)=\lambda_j I_{n_j}+[N_{1j}, N_{2j}, \cdots, N_{s_j,j}]=\lambda_j I_{nj}+N_{nj},$$

易知 N_{nj} 也是幂零阵，故

$$A=P[\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \cdots, \lambda_s I_{n_s}]P^{-1}$$

$$+P[N_{n_1}, N_{n_2}, \cdots, N_{n_s}]P^{-1}=S+N,$$

显然，$N=P[N_{n_1}, N_{n_2}, \cdots, N_{n_s}]P^{-1}$ 是幂零阵，且易证 $SN=NS$。

证毕。

Jordan 标准形在“矩阵方程论”与“矩阵函数论”中都有它的应用，而这些内容在后继课程“常微分方程”以及“现代控制论”中都是必需的。有关 Jordan 标准形的各种主要应用将见之于本节习题及本章选做题中。

最后说明一点：有关方阵的相似标准形，除了本章所说的三种主要标准形，即有理标准形、Jacobson 标准形（均对任意数域来说）、Jordan 标准形(对复数域来说)外，还可作出其他标准形。事实上，只要能作出相应于A的初等因子组的各种矩阵块，则这些矩阵块所成的分块对角阵就有可能成为A的相似标准形，这个想法是重要的。

习　　题

1．求下列方阵的 Jordan 标准形

(i) $\begin{pmatrix} \sqrt{-1} & \sqrt{-1} \\ 1 & \sqrt{-1} \end{pmatrix}$；　　(ii) $\begin{pmatrix} 1 & 1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix}$；

(iii) $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$；　　(iv) $\begin{pmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{pmatrix}$.

2．设 $A^l = I_n$，l 是正整数，求证：A 与对角阵 D 相似，且 D 的主对角元的模是 1.

3．求证下列各对 n 阶复方阵相似

(i) $A=\begin{pmatrix} 0 & 1 \\ I_{n-1} & 0 \end{pmatrix}$，　　$B=[\zeta_1, \zeta_2, \cdots\cdots, \zeta_n]$.

这里 $\zeta_1, \zeta_2, \cdots, \zeta_n$ 是 1 的不同的 n 次根(即 $\zeta_i^n=1$, $i=1, 2, \cdots, n$)，

(ii) $A=\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}$，$B=\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & * & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}$，

B 的 * 处的元素可以是任意数.

4．设 λ_1 是 n 阶复方阵 A 的 k 重特征值，求证：

$$r((\lambda_1 I_n - A)^k) = n-k.$$

5．设 $J_s(\lambda_i)$ 是 s 阶 Jordan 块，$\lambda_i \neq 0$，求 $J_s(\lambda_i)^{-1}$.

6．设 3 阶阵 A 的 Jordan 标准形是：

$$J=\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

且使 $P^{-1}AP=J$ 的 P 是：

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix},$$

如果 $f(\lambda)=\lambda^3-4\lambda^2+2$, 试求 $f(A)$。

7. 设 n 阶非异阵 P 使

$$P^{-1}AP=\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix},$$

把 P 按它的列分块: $P=(\delta_1, \delta_2, \cdots, \delta_n)$, 试求 $\delta_1, \delta_2, \cdots, \delta_n$ 所满足的方程组。

8. 求秩为1的 n 阶复方阵的 Jordan 标准形。

9. 设 $r(A)=tr(A)=1$, 求证: A 必是幂等阵。

10. 设 n 阶阵 A 的特征多项式是 $(\lambda-1)^n$, 对任一正整数 l, $2\leqslant l\leqslant n$, 求证: A^l 与 A 相似。

选 做 题

1. 设 $A(\lambda)$、$B(\lambda)$ 都是 $m\times n$ λ-阵, 它的元素是数域 K 上的多项式, 如果存在 m 阶非异 λ-阵 $P(\lambda)$ 与 n 阶非异 λ-阵 $Q(\lambda)$, 使 $P(\lambda)A(\lambda)Q(\lambda)=B(\lambda)$, 则称 $A(\lambda)$ 与 $B(\lambda)$ 相抵, 求证:

(i) 任何 $m\times n$ λ-阵必与 $m\times n$ 对角形 λ-阵:

$$\begin{pmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

相抵, 其中 $d_r(\lambda)\neq 0$, $d_i(\lambda)\mid d_{i+1}(\lambda)$, 且 $d_r(\lambda)$, $d_i(\lambda)$ 都是首一多项式, $i=1, 2, \cdots\cdots, r-1$, 称这个 r 为原来的 $m\times n$ λ-阵的秩。

(ii) $m\times n$ λ-阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵的充要条件是, 它们有相同的秩与相同的不变因子(组)(其定义与 n 阶 λ-阵的不变因子的定义相同)。

(iii) $m\times n$ λ-阵 $A(\lambda)$ 与 $B(\lambda)$ 相抵的充要条件是, 它们有相同的秩与

相同的初等因子组(其定义与 n 阶 λ-阵的初等因子组的定义相同)。

2. 设 A 与 B 是数域 K 上的同阶方阵，求证：$\lambda I-A$ 与 $\lambda I-B$ 相抵的充要条件是，A 与 B 有分解式：

$$A=PQ,\quad B=QP,$$

其中 P 与 Q 都是 K 上的方阵，且其中至少有一个是非异阵。

3. 设 A_1、A_2、B_1、B_2 都是同阶方阵，且 A_1 与 A_2 都是非异阵，求证：λA_1-B_1 与 λA_2-B_2 相抵的充要条件是，存在非异阵 P 与 Q，使得

$$PA_1Q=A_2,\ PB_1Q=B_2.$$

4. 设 A 是数域 K 上的 n 阶非异阵，B 是 K 上的 n 阶对合阵，如果 $\lambda B-A$ 与 $\lambda B-A^{-1}$ 相抵，求证：A 必可分解为 K 上三个对合阵的乘积。

5. 设 A 是有理数域 Q 上的正交阵（即 A 的元素都是有理数，且 $AA'=I_n$）。求证：

(i) 矩阵方程 $AXA=X$ 在 Q 上必有非异阵解：$X=P$(P 是非异阵)。

(ii) A 必可分解为 Q 上两个对合阵的乘积。

(注意：下一章将证明，实数域 R 上的正交阵可以分解为 R 上两个对称、对合阵的乘积。)

6. 如果数域 K 上的 $2n$ 阶阵 A 满足：

$$A\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix}A'=\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix},$$

则称 A 为(K 上的)辛矩阵。求证：

(i) 若 A 是 K 上的辛矩阵，则矩阵方程 $AXA=X$ 在 K 上必有非异阵解。

(ii) K 上的辛矩阵必可分解为 K 上两个对合阵的乘积。

7. 求秩为 1 的 $n(n>2)$ 阶阵 A 的第 $n-1$ 个行列式因子，并求 λI_n-A 的法式。

8. 设 A 是数域 K 上的 n 阶阵，则矩阵方程 $AX-XA'=0$ 在 K 上必有非异对称阵解。(提示：把 A' 写成它的 Frobenius 标准形，

$$A=Q[F_1',F_2',\cdots,F_{n-k}']Q^{-1},$$

再把 $AX-XA'=0$ 写成同解方程：

$$[F_1',F_2',\cdots,F_{n-k}'](Q^{-1}XQ'^{-1})$$
$$-(Q^{-1}XQ'^{-1})[F_1,F_2,\cdots,F_{n-k}]=0,$$

然后设 $Q^{-1}XQ=[X_{11},X_{22},\cdots,X_{n-k,n-k}]$，再把上矩阵方程化为：

$$F_i'X_{ii}-X_{ii}F_i=0;\qquad i=1,2,\cdots,n-k,$$

即可解出对称阵解$[X_{11}, \cdots, X_{n-k,n-k}]$)。

9．求证数域K上任何方阵必可分解为K上两个对称阵（即满足$A'=A$的阵)的乘积。(提示：应用第8题)

10．设F是A的有理标准形的一个r阶非异Frobenius块：

$$F=\begin{pmatrix} 0 & -a_r \\ I_{r-1} & \alpha \end{pmatrix}, \alpha=(-a_{r-1}, -a_{r-2}, \cdots, -a_1)', a_r \neq 0,$$

(i) 求证：$F=CQD$，其中，

$$C=\begin{pmatrix} -1 & 0 \\ a_r^{-1}\alpha & I_{r-1} \end{pmatrix}, \quad Q=\begin{pmatrix} 0 & \pm 1 \\ I_{r-1} & 0 \end{pmatrix}, \quad D=\begin{pmatrix} I_{r-1} & 0 \\ 0 & \pm a_r \end{pmatrix}.$$

(ii) 如果数域K上对角阵D的主对角元两两互为逆元(当D是奇数阶阵时，则主对角元中有奇数个± 1)，求证：D可分解为K上的一个正交阵(即方阵的转置阵与它的逆阵相等)与一个对合阵的乘积。

(iii) 设t是任一非零数，则F必有分解式：

$$F=TC_1Q_1D_1T^{-1},$$

而C_1、Q_1、D_1分别是下面的对合阵、正交阵、对角阵，并且

$$T=\begin{pmatrix} t^{-1} & 0 \\ 0 & I_{r-1} \end{pmatrix}.$$

(iv) (Gustafson-Halmos-Radjavi, 1976) 如果A是数域K上的幺模阵(即$|A|=\pm 1$)，应用A的有理标准形及(i)、(ii)、(iii)证明下面的结论：

“数域K上的任何幺模阵必可分解为K上个数不超过4个的对合阵的乘积。”

11．如果数域K上的n阶阵A的不变因子(组)是

$$1, \cdots, 1, f(\lambda)=|\lambda I_n - A|,$$

且$AB=BA$，求证：$B=g(A)$，其中$g(\lambda)$是K上的某一多项式。

12．设A、B、C分别是数域K上的m阶阵、n阶阵、$m\times n$阵，B的非常数不变因子是$d_{k+1}(\lambda)$，$d_{k+2}(\lambda)$，$\cdots$，$d_n(\lambda)$，且

$$P^{-1}BP=\begin{pmatrix} F_1 & & & \\ & F_2 & & \\ & & \ddots & \\ & & & F_{n-k} \end{pmatrix},$$

其中P是K上非异阵，F_i是$d_i(\lambda)$的友阵(Frobenius块)，$i=1, 2, \cdots, n$。

试写出矩阵方程 $AX+XB'=C$ 在K上有解的充要条件。

13. 求证：数域K上n阶阵A相似于对角阵的充要条件是：$r(\lambda I_n-A)=r(\lambda I_n-A)^2)$。其中$\lambda$是$A$(在$K$上)的任一特征值。
(提示：先应用第三章选做题第4题的结论，再应用本章§4的习题第5题的结论。)

14. 设

$$J_s(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & & \\ & \lambda_i & 1 \ddots & \\ & & \ddots & 1\\ & & & \lambda_i\end{pmatrix}$$

是n阶阵A的s阶 Jordan 块，$f(\lambda)$是k次多项式，求证：

$$f(J_s(\lambda_i))=\begin{pmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \frac{f''(\lambda_i)}{2!} & \cdots & \frac{f^{(s-1)}(\lambda_i)}{(s-1)!}\\ & f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \ddots & \vdots\\ & & \ddots & \ddots & \frac{f'(\lambda_i)}{1!}\\ & & & & f(\lambda_i)\end{pmatrix},$$

並求$f(A)$。

15. 设n阶阵A的不同特征值为$\lambda_1, \lambda_2, \cdots, \lambda_t$，又设

$$P^{-1}AP=[J(\lambda_1), J(\lambda_2), \cdots, J(\lambda_t)],$$

其中

$$J(\lambda_j)=[J_{l_{1j}}(\lambda_j), J_{l_{2j}}(\lambda_j), \cdots, J_{l_{s_j j}}(\lambda_j)],\ j=1, 2, \cdots, t,$$

而$J_{l_{ij}}(\lambda_j)$是l_{ij}阶 Jordan 块，又设A的最小多项式是

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_t)^{l_t},$$

$g(\lambda)$ 是有限 l 次可导函数，$l\geqslant\max\{l_1, l_2, \cdots, l_t\}$，如果$g^{(k)}(\lambda_j)$存在，$k=0, 1, \cdots, l_j-1, j=1, 2, \cdots, t$，则记

$$g(A)=P[g(J(\lambda_1)), g(J(\lambda_2)), \cdots, g(J(\lambda_t))]P^{-1},$$

其中

$$g(J(\lambda_j))=[g(J_{l_{1j}}(\lambda_j)), g(J_{l_{2j}}(\lambda_j)), \cdots, g(J_{l_{s_j j}}(\lambda_j))],$$

而

$$g(J_{l_{ij}}(\lambda_j))=\begin{pmatrix} g(\lambda_j) & \frac{g'(\lambda_j)}{1!} & \cdots\cdots & \frac{g^{(l_{ij}-1)}(\lambda_j)}{(l_{ij}-1)!} \\ & g(\lambda_j) & \ddots & \vdots \\ & & \ddots & \frac{g'(\lambda_j)}{1!} \\ & & & g(\lambda_j) \end{pmatrix}$$

称 $g(A)$ 为矩阵 A 的函数值或**矩阵函数**。

(i) 设 $g_1(\lambda)$, $g_2(\lambda)$ 都是 l 次可导函数，且 $g(\lambda)=g_1(\lambda)+g_2(\lambda)$，求证：$g(A)=g_1(A)+g_2(A)$。

(ii) $g_1(\lambda)$, $g_2(\lambda)$ 的假设同(i)，且 $\varphi(\lambda)=g_1(\lambda)g_2(\lambda)$，求证：$\varphi(A)=g_1(A)g_2(A)$。

(iii) 如果非异阵 Q 使 $A=Q[J(\lambda_1), J(\lambda_2), \cdots, J(\lambda_t)]Q^{-1}$, $Q\neq P$，求证：

$$g(A)=Q[g(J(\lambda_1)), g(J(\lambda_2)), \cdots, g(J(\lambda_t))]Q^{-1},$$

即 $g(A)$ 的定义与变换矩阵(P 与 Q)的取法无关。

(提示：用插值公式，即存在 $\sum_{i=1}^{t} l_i-1$ 次多项式 $\xi(\lambda)$，使得

$$\xi^{(k)}(\lambda_j)=g^{(k)}(\lambda_j),\ k=0, 1, \cdots, l_j;\ j=1, 2, \cdots, t)。$$

16. 已知 A 的 Jordan 标准形 J 及变换矩阵 $P(P^{-1}AP=J)$ 试决定与 A 可交换的阵 B 的形状。

17. 设

$$R=\begin{pmatrix} r_{11} & & * \\ & r_{22} & \\ & & \ddots \\ & & & r_{nn} \end{pmatrix}$$

是上三角阵，C 是 n 阶阵，求证：

(i) 矩阵方程 $RX+XR'=C$ 有唯一解的充要条件是：

$$r_{ii}+r_{jj}\neq 0,\ i, j=1, 2, \cdots, n。$$

(ii) 当 $r_{ii}+r_{jj}\neq 0$, $i, j=1, 2, \cdots\cdots, n$, R 是实方阵，且 C 是实对称阵时，则 $RX+XR'=C$ 的解 X 必是实对称阵。

18. 用方阵的 Jordan 标准形证明：矩阵方程 $AX+XB'=C$ 有唯一解的充要条件是：$\lambda_i+\mu_j\neq 0$, $i=1, 2, \cdots, m$; $j=1, 2, \cdots, n$。此处 A、B、C 分别是 m 阶阵、n 阶阵、$m\times n$ 阶阵，而 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是 A 的特征值，μ_1,

$\mu_2, \cdots, \mu_n$ 是 B 的特征值。

(提示：仿 19 题的证法。)

19. 设 A、B、C 分别是 m 阶阵、n 阶阵、$m\times n$ 阵。

(i) 记 $g(\lambda)=|\lambda I_n+B|$，求证：矩阵方程 $AX+XB'=C$ 有唯一解的充要条件是，$g(A)$ 为非异阵。

(ii) 当 $m=n$，$A=B$ 时，记

$$f(\lambda)=\frac{1}{2}(|\lambda I_n+A|+|\lambda I_n-A|),$$

$$h(\lambda)=\frac{1}{2}(|\lambda I_n+A|-|\lambda I_n-A|),$$

求证：矩阵方程 $AX+XA'=C$ 有解的充要条件是，$f(A)$（或 $h(A)$）为非异阵。

(提示：把 B 化为它的 Jordan 标准形，並仿照定理 3.4 的证法可证得 (i)。对 (ii)，由 (i) 及 Hamilton-Cayley 定理即可证得 (ii)。)

20. 应用"矩阵方阵 $AX-XB'=0$ 有非零解的充要条件是，$f(B)$ 为奇异阵"这一结论（见 21 题 (i)）重新证明 Barnett 定理：m 次多项式 $f(\lambda)$ 与 n 次多项式 $g(\lambda)$ 有公根的充要条件是，$f(B)$ 为奇异阵，其中 B 是 $g(\lambda)$ 的友阵（参阅第四章 §7）。

(提示：设 $f(\lambda)$ 的友阵为 A，考虑 $AX-XB'=0$ 有非零解的充要条件，並应用第 20 题、21 题 (i) 的结论。)

第七章　镜象阵、方阵的正交相似与酉相似

本章介绍镜象阵(镜面反射阵)的基本概念，并用镜象阵处理了方阵正交相似与酉相似的各种问题，讨论了它们的各种应用。

§1　镜象阵的概念、基本定理

一、镜象阵的概念

镜象阵，也称 Hauseholder 初等反射阵，或镜面反射阵，这一概念是由物体在平面镜中成象的一些简单物理性质发展而来，现在就来说明这一点。

在通常的三维几何空间中，任一点 S(向量 α)对平面(镜面) π 的象是镜面下的一点 S'(向量 β)，如右边图 1 所示，根据成象的物理意义，它有如下性质：

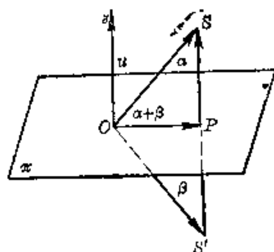

图 1

(i) α 与 β 的长度相等，即 $(\alpha'\alpha)^{1/2}=(\beta'\beta)^{1/2}$($\alpha$ 与 β 都是三维列向量)；

(ii) 点 S' 在由过 S 而与 π 垂直的直线上，对 π 来说，S 与 S' 具有对称的位置，即 S 可以看作 S' 的镜象，由原点 O 作 π 的单位法向量 u (即 $u'u=1$)，则 $\overline{Oz}$ 平行于 $\overline{SS'}$，所以

$$\beta-\alpha=ku=uk \quad (k \text{ 为实数}). \tag{1}$$

今求 k。在(1)式的两边左乘 u'，则得

$$u'(\beta-\alpha)=(u'u)k=k, \tag{2}$$

又因为 $\alpha+\beta$ 位于 π 上，所以

$$u'(\alpha+\beta)=0, \tag{3}$$

(3)式减去(2)式即得 $k=-2u'\alpha$，把它代入(1)的最后一个等式中得到：

$$\beta-\alpha=-2uu'\alpha,$$

于是

$$\beta=(I_3-2uu')\alpha, \tag{4}$$

故 α 与它的(镜)象 β 可以通过 (4) 式中的 3 阶阵 I_3-2uu' 联系起来，所以称 I_3-2uu' 为 3 阶实镜象阵。一般引进

定义 设 u 是 n 维实的单位列向量(即 $u'u=1$)，则称

$$H=I_n-2uu'$$

为(n 阶)**实镜象阵**，而对 n 维复的单位列向量 u (即 $\bar{u}'u=1$)，则称 $H^*=I_n-2u\bar{u}'$ 为(n 阶)**复镜象阵**。

命题 1 如果 H 是 n 阶实镜象阵，则 $[I_m, H]$ 也是实镜象阵，

【证明】 因为 $H=I_n-2uu'$，而 $u'u=1$，所以

$$[I_m, H]=\begin{pmatrix}I_m & 0\\ 0 & I_n-2uu'\end{pmatrix}=\begin{pmatrix}I_m & 0\\ 0 & I_n\end{pmatrix}-2\begin{pmatrix}0 & 0\\ 0 & uu'\end{pmatrix}$$

$$=I_{m+n}-2\begin{pmatrix}0\\ u\end{pmatrix}(0, u'),$$

而 $\begin{pmatrix}0\\ u\end{pmatrix}$ 显然是 $m+n$ 维单位列向量，故 $[I_m, H]$ 是 ($m+n$ 阶)实镜象阵。 证毕。

如把命题 1 中的 H 换成复镜象阵 H^*，则相应的结论仍然成立。

命题 2 H 是正交、对称阵（因而也是实对合阵）。H^* 是 Hermite、酉阵。(因而也是复对合阵。)

【证明】 因为

$$\overline{(H^*)}'=\overline{(I_n-2u\bar{u}')}'=I_n-2u\bar{u}'=H^*,$$

$$\overline{(H^*)}'\cdot H^*=I_n+4u(\bar{u}'u)\bar{u}'-4u\bar{u}'=I_n(\text{因 } \bar{u}'u=1),$$

故 H^* 是 Hermite、酉阵，当 u 取实向量，即得本命题的前一个结论。 证毕。

命题 3 $|H|=-1$, $|H^*|=-1$。

这在第二章中已用行列式的第二降阶定理证明过了。

二、基本定理

对三维空间中任意两个长度相等的向量 α 与 β，可把它们看作过原点而其端点分别在 S 与 S' 上的两个向量，作过原点而垂直于线段 $\overline{SS'}$ 的二等分线的平面 π（见图 1），则 β 可看作 α 的镜象，于是由上一段的讨论可知，存在 3 阶实镜象阵 H，把 α "变成" β，即 $H\alpha=\beta$，而由命题 2 可得，$H\beta=\alpha$，也即 α 可看作 β 的镜象，那么，对"长度"相等的两个 n 维实的列向量 α 与 β，是否也存在 n 阶实镜象阵 H，使 $H\alpha=\beta$ 呢？回答是肯定的，为此先要引进 n 维列向量的长度的概念。

定义 设 α 是 n 维复的列向量，称 $\|\alpha\|=\sqrt{\overline{\alpha}'\alpha}$ 为 α 的**长度**。

当 α 是实向量时，则 $\|\alpha\|=(\alpha'\alpha)^{1/2}$。今证镜象阵理论中的基本定理。

定理 1.1 设 α,β 是两个不同的 n 维实的列向量，如果

$$\|\alpha\|=\|\beta\|,$$

则必存在 n 阶实镜象阵 H，使 $H\alpha=\beta$。

【证明】 由假设 $\alpha\neq\beta$，故 $\alpha-\beta\neq0$，于是

$$\|\alpha-\beta\|=\sqrt{\overline{(\alpha-\beta)}'(\alpha-\beta)}\neq0,$$

作向量

$$u=\frac{\alpha-\beta}{\|\alpha-\beta\|},\tag{5}$$

则易知 u 是单位列向量，即 $u'u=1$。把(5)式改写为

$$\beta-\alpha=-u\|\alpha-\beta\|。\tag{6}$$

另一方面，由假设 $\|\alpha\|=\|\beta\|$，故 $\alpha'\alpha=\beta'\beta$，于是

$$\begin{aligned}\|\alpha-\beta\|^2&=(\alpha-\beta)'(\alpha-\beta)=\alpha'\alpha+\beta'\beta-\beta'\alpha-\alpha'\beta\\&=2\alpha'\alpha-2\beta'\alpha=2(\alpha-\beta)'\alpha。\end{aligned}\tag{7}$$

把(7)式改写成

$$\|\alpha-\beta\|=2\left(\frac{\alpha-\beta}{\|\alpha-\beta\|}\right)'\alpha=2u'\alpha,\tag{8}$$

将(8)式代入(6)式，经整理后即得

$$\beta=\alpha-2uu'\alpha=(I_n-2uu')\alpha=H\alpha.$$

证毕。

上述证明是构造性的。由 α 与 β 可以求出 u，从而找出 H。

例 1 设 $\alpha=(3,4)'$，$\beta=(5,0)'$，求 2 阶实镜象阵 H，使

$$H\alpha=\beta.$$

【解】 因为 $\alpha-\beta=(-2,4)'$，$\|\alpha-\beta\|=2\sqrt{5}$，所以

$$u=\frac{\alpha-\beta}{\|\alpha-\beta\|}=\begin{pmatrix}-1/\sqrt{5}\\ 2/\sqrt{5}\end{pmatrix},$$

于是

$$H=I_2-2uu'=\begin{pmatrix}3/5 & 4/5\\ 4/5 & -3/5\end{pmatrix}$$

容易验证 $H\alpha=\beta$。

推论 1 设 α 是 n 维实的列向量，$\alpha\neq 0$，u 是 n 维实的单位列向量，则必存在实镜象阵 H，使

$$H\frac{\alpha}{\|\alpha\|}=u. \tag{9}$$

【证明】 因为 $\frac{\alpha}{\|\alpha\|}$ 也是单位列向量，与 u 的长度相等，故由定理 1.1 即得本推论。 证毕。

推论 2 设 α,β 是任意两个不同的 n 维实的列向量，则必存在 n 阶实镜象阵 H，使

$$H\frac{\alpha}{\|\alpha\|}=\frac{\beta}{\|\beta\|},$$

这由定理 1.1 可知，它是显然的。

对 n 维复的列向量，有定理 1.1 的如下推广结果：

定理 1.2 设 α，β 是两个 n 维复的列向量，$\alpha\neq\beta$，如果 $\bar{\alpha}'\beta$ 是实数，则必存在 n 阶复镜象阵 H^*，使 $H^*\alpha=\beta$。

【证明】 作单位向量：

$$u=\frac{\alpha-\beta}{\|\alpha-\beta\|},$$

则

$$\beta-\alpha=-u\|\alpha-\beta\|. \tag{10}$$

但是

$$\|\alpha-\beta\|^2=\overline{(\alpha-\beta)}'(\alpha-\beta)=\bar{\alpha}'\alpha+\bar{\beta}'\beta-\bar{\beta}'\alpha-\bar{\alpha}'\beta. \tag{11}$$

由假设，$\bar{\alpha}'\alpha=\bar{\beta}'\beta$，$\bar{\alpha}'\beta=\overline{(\bar{\alpha}'\beta)'}=\bar{\beta}'\alpha$，故(11)式可写成

$$\|\alpha-\beta\|^2=2\overline{(\alpha-\beta)}'\alpha,$$

也即

$$\|\alpha-\beta\|=2\overline{\left(\frac{\alpha-\beta}{\|\alpha-\beta\|}\right)}'\alpha=2\bar{u}'\alpha.$$

把上式代入(10)式即得

$$\beta=(I_n-2u\bar{u}')\alpha=H^*\alpha.$$ 证毕.

例 2 设 $\alpha=(1-\sqrt{-1},1+\sqrt{-1})'$，$\beta=(\sqrt{2},\sqrt{2})'$ 求 2 阶复镜象阵 H^*，使 $H^*\alpha=\beta$.

【解】 因为 $\|\alpha\|=\|\beta\|=2$，并且

$$\bar{\alpha}'\beta=(1+\sqrt{-1})\sqrt{2}+(1-\sqrt{-1})\sqrt{2}=2\sqrt{2}$$

是实数，故满足定理 1.2 的条件，作单位向量：

$$u=\frac{\alpha-\beta}{\|\alpha-\beta\|}$$

$$=\frac{1}{2\sqrt{2-\sqrt{2}}}(1-\sqrt{2}-\sqrt{-1},1-\sqrt{2}+\sqrt{-1})',$$

于是

$$H^*=I_2-2uu'=\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{2+\sqrt{2}}{2}$$

$$\cdot\begin{pmatrix}2-\sqrt{2}&(1-\sqrt{2})(1-\sqrt{-1})\\(1-\sqrt{2})(1+\sqrt{-1})&2-\sqrt{2}\end{pmatrix}$$

$$=\begin{pmatrix}0&\frac{\sqrt{2}}{2}(1-\sqrt{-1})\\\frac{\sqrt{2}}{2}(1+\sqrt{-1})&0\end{pmatrix},$$

容易验证：$H\alpha=\beta$.

三、任意方阵的 QR 分解

定理 1.3 设 A 是任意 n 阶实方阵，则 A 必有分解式：

$$A=QR, \tag{12}$$

其中 Q 为正交阵，R 为主对角元全是非负数的上三角阵。称（12）式为 A 的 QR 分解。

【证明】 把 A 按它的列分块：

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}=(\alpha_1,\alpha_2,\cdots,\alpha_n).$$

当 $n=1$ 时，因为

$$a_{11}=\begin{cases} 1\cdot a_{11}, & \text{当 } a_{11}\geqslant 0, \\ (-1)(-a_{11}), & \text{当 } a_{11}<0, \end{cases}$$

故定理对 $n=1$ 是成立的。设定理对 $n-1$ 成立，今证定理对 n 也成立。

(i) 如果 $\alpha_1=0$，则

$$A=\begin{pmatrix} 0 & * \\ 0 & A_1 \end{pmatrix},$$

其中 A_1 是 $n-1$ 阶阵，故由归纳法假设，

$$A_1=Q_1R_1,$$

其中 Q_1 是 $n-1$ 阶正交阵，R_1 是上三角阵，且它的主对角元全是非负数，于是

$$A=\begin{pmatrix} 0 & * \\ 0 & A_1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}\begin{pmatrix} 0 & * \\ 0 & R_1 \end{pmatrix}=QR,$$

易知

$$Q=\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}, \qquad R=\begin{pmatrix} 0 & * \\ 0 & R_1 \end{pmatrix}$$

分别是正交阵与主对角元都是非负数的上三角阵。

(ii) 当 $\alpha_1\neq 0$，则令

$$\beta_1=(\|\alpha_1\|,0,\cdots,0)',$$

显然 $\|\beta_1\|=\|\alpha_1\|$，故由定理 1.1，存在实镜象阵 H，使 $H\alpha_1=\beta_1$，所以

$$HA=(H\alpha_1,H\alpha_2,\cdots,H\alpha_n)=(\beta_1,H\alpha_2,\cdots,H\alpha_n)=\begin{pmatrix}\|\alpha_1\| & * \\ 0 & A_1\end{pmatrix}, \tag{13}$$

且 $\|\alpha_1\|>0, A_1$ 是 $n-1$ 阶阵，由归纳法假设，$A_1=Q_1R_1$，其中 Q_1 与 R_1 分别是 $n-1$ 阶正交阵与主对角元全是非负数的上三角阵。又因 $H^{-1}=H$（命题 2），于是(13)式化为

$$A=H\begin{pmatrix}\|\alpha_1\| & * \\ 0 & Q_1R_1\end{pmatrix}=H\begin{pmatrix}1 & 0 \\ 0 & Q_1\end{pmatrix}\begin{pmatrix}\|\alpha_1\| & * \\ 0 & R_1\end{pmatrix}=QR,$$

易知

$$Q=H\begin{pmatrix}1 & 0 \\ 0 & Q_1\end{pmatrix} \qquad R=\begin{pmatrix}\|\alpha_1\| & 0 \\ 0 & R_1\end{pmatrix}$$

分别是正交阵与主对角元全是非负数的上三角阵。　　　　证毕。

不难看出，定理 1.3 的证明是构造性的。

例 3　求

$$A=\begin{pmatrix}2 & -1 & 3 \\ 0 & 3 & -2 \\ 0 & 4 & 1\end{pmatrix}$$

的 QR 分解。

【解】　因为方阵 $\begin{pmatrix}3 & -2 \\ 4 & 1\end{pmatrix}$的第一列 α_1 的长度是 5，令

$$\beta_1=(\|\alpha_1\|,0)'=(5,0)',$$

则 $\|\alpha_1\|=\|\beta_1\|$，而

$$u=\frac{\alpha_1-\beta_1}{\|\alpha_1-\beta_1\|}=\begin{pmatrix}-1/\sqrt{5} \\ 2/\sqrt{5}\end{pmatrix},$$

所以

$$H_1=I_2-2uu'=\begin{pmatrix}3/5 & 4/5 \\ 4/5 & -3/5\end{pmatrix},$$

于是由 $H_1^2=I_2$ 可得

$$A=\begin{pmatrix}1 & 0\\0 & H_1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & H_1\end{pmatrix}\begin{pmatrix}2 & -1 & 3\\0 & 3 & 2\\0 & 4 & 1\end{pmatrix}$$

$$=\begin{pmatrix}1 & 0 & 0\\0 & 3/5 & 4/5\\0 & 4/5 & -3/5\end{pmatrix}\begin{pmatrix}2 & -1 & 3\\0 & 5 & -2/5\\0 & 0 & -11/5\end{pmatrix}$$

$$=\begin{pmatrix}1 & 0 & 0\\0 & 3/5 & 4/5\\0 & 4/5 & -3/5\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{pmatrix}$$

$$\cdot\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{pmatrix}\begin{pmatrix}2 & -1 & 3\\0 & 5 & -2/5\\0 & 0 & -11/5\end{pmatrix}$$

$$=\begin{pmatrix}1 & 0 & 0\\0 & 3/5 & -4/5\\0 & 4/5 & 3/5\end{pmatrix}\begin{pmatrix}2 & -1 & 3\\0 & 5 & -2/5\\0 & 0 & 11/5\end{pmatrix}.$$

上面最后一个等式的两个方阵的乘积即为所求的A的QR分解。

当A是非异阵时，A的QR分解中的R显然也是非异阵，故得

推论 1 设A是非异阵，则$A=QR$，其中Q是正交阵，R是主对角元全是正数的上三角阵。并且这种分解是唯一的。

【证明】 本推论的第一部分已显然可得。今证分解的唯一性，设

$$A=QR=Q_1R_1 \tag{14}$$

是A的任意两种QR分解，其中Q、Q_1都是正交阵，R、R_1都是主对角元全大于零的(非异)上三角阵，于是(14)式可改写为

$$Q_1^{-1}Q=R_1R^{-1}. \tag{15}$$

由于$Q_1^{-1}Q$仍是正交阵，故R_1R^{-1}是正交阵、又是主对角元全大于零的上三角阵，而正交的上三角阵只能是对角阵，且其主对角元是1或-1(这是容易证明的)，但R_1R^{-1}的主对角元全是正数，故R_1R^{-1}只能是单位阵：$R_1R^{-1}=I_n$，故$R_1=R$，再由(15)式得到

$$Q=Q_1.$$

证毕。

推论 2 任何正交阵A必可分解为有限个实镜象阵的乘积。

【证明】 由定理1.3的证明过程容易看出，必存在有限个实镜象阵H_1、H_2、…、$H_s(s\leqslant n)$，使

$$H_s\cdots H_2H_1A=\begin{pmatrix}\|\alpha\| & & & \\ & \|\delta\| & & * \\ & & \ddots & \\ & & & \|\eta\|\end{pmatrix}=R, \tag{16}$$

此处R是主对角元全大于零的上三角阵。

由假设A是正交阵，而$H_s\cdots H_2H_1$仍是正交阵，故由(16)式，R是主对角元全是正数的正交上三角阵，于是由推论1证明中的说明可知，$R=I_n$。所以(16)式化为

$$H_s\cdots H_2H_1A=I_n,$$

由$H_i^2=I_n, i=1,2,\cdots,s$即得

$$A=H_1H_2\cdots H_s. \qquad\qquad \text{证毕。}$$

因为有限个实镜象阵的乘积自然是正交阵，这个结论与推论2合起来，即得

推论 3 A是正交阵的充要条件是，它可分解为有限个实镜象阵的乘积。

由于镜象阵有良好的性质，且容易找到，故由推论3可知，要找满足某种条件的正交阵，只要找满足某种条件的有限个实镜象阵就可以了，以下将会反复运用这一有用的想法。

定理 1.4 任何复方阵A必可分解为一个酉阵U与一个上三角阵R的乘积：

$$A=UR. \tag{17}$$

【证明】 把A按它的列分块：

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}=(\alpha_1,\alpha_2,\cdots,\alpha_n).$$

(i) 如果 $a_{11}\neq 0$，把复数 a_{11} 写成

$$a_{11}=a_1+b_1\sqrt{-1}=|a_{11}|e^{\sqrt{-1}\varphi_1}\quad (a, b \text{ 是实数}),$$

此处 $|a_{11}|=\sqrt{a_1^2+b_1^2}$，$\cos\varphi_1=\dfrac{a_{11}}{|a_{11}|}$. 令

$$\beta_1=(\|\alpha_1\|e^{\sqrt{-1}\varphi}, 0, \cdots, 0)',$$

则 $\|\alpha_1\|=\|\beta_1\|$，并且

$$\bar{\alpha}_1'\beta_1=|a_{11}|\cdot\|\alpha_1\|$$

是实数，故由定理 1.2，必存在复镜象阵 H_1^*，使 $H_1^*\alpha_1=\beta_1$. 于是

$$H_1^*A=(\beta_1, H_1^*\alpha_2, \cdots, H_1^*\alpha_n)=\begin{pmatrix}\|\alpha_1\|e^{\sqrt{-1}\varphi} & * \\ 0 & A_1\end{pmatrix}.$$

至此，以下的证明与实方阵的证明相同，故不再重复。

(ii) 如果 $a_{11}=0$，则取

$$\beta_1=(\|\alpha_1\|, 0, \cdots, 0)'$$

这时 $\|\alpha_1\|=\|\beta_1\|$，并且 $\bar{\alpha}_1'\beta_1=0$（自然是实数），所以同 (i)，可找到复镜象阵 M_1，使

$$M_1A=\begin{pmatrix}\|\alpha_1\| & * \\ 0 & A_1\end{pmatrix},$$

于是用归纳法即可证得本定理。　　　　证毕。

称(17)式为 A 的 **UR 分解**。

习　题

1. 找实镜象阵 H，使 $H\alpha=\beta$.

(i) $\alpha=(0, 1)'$，$\beta=(1, 0)'$；

(ii) $\alpha=(1, 1, 0)'$，$\beta=\left(\dfrac{11}{9}, \dfrac{5}{9}, \dfrac{4}{9}\right)'$；

(iii) $\alpha=e_1=(1, 0, 0, 0)'$，$\beta=\left(\dfrac{1}{2}, -\dfrac{1}{2}, \dfrac{1}{2}, -\dfrac{1}{2}\right)'$.

2. 找复镜象阵 H^*，使 $H^*\alpha=\beta$，此处 $\alpha=(-\sqrt{-2}, \sqrt{-2})$，

$$\beta=(\sqrt{2}, \sqrt{2})'.$$

3. 求下列方阵的 QR 分解(UR 分解)

(i) $\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$, (ii) $\begin{pmatrix} \sqrt{-1} & 2 \\ 3 & 1-\sqrt{-1} \end{pmatrix}$, (iii) $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 3 \\ 5 & 2 & 5 \end{pmatrix}$.

4. 求证：正交阵的行列式等于 1 (或 −1)的充要条件是，A 可分解为偶数个(或奇数个)实镜象阵的乘积。

5. 求证：任何非异实方阵 A 必有唯一分解式：

$$A = QL,$$

其中 Q 是正交阵，L 是主对角元全是正数的下三角阵。

. 求证：任何非异实方阵必有唯一分解式：

$$A = LQ = RQ_1,$$

此处 Q、Q_1 都是正交阵，L 与 R 分别是主对角元全是正数的下三角阵与上三角阵。

§2 实方阵正交相似的矩阵

定义 如果 A 与 B 都是实方阵，且存在正交阵 Q，使

$$Q^{-1}AQ = B \quad (\text{即} \ Q'AQ = B), \tag{1}$$

则称 A 与 B **正交相似**，或称 A 正交相似于 B。

正交相似比一般的相似概念要求强些，然而却在不少问题中遇到它，本节讨论任意一个实方阵能够正交相似于何种形状简单的矩阵。

一、实方阵正交相似于对角阵的条件

定理 2.1 设 $A = (a_{ij})_{n\times n}$ 是实方阵，如果 A 的特征值都是实数，则必可找到正交阵 Q，使

$$Q'AQ = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ & 0 & & \lambda_n \end{pmatrix}, \tag{2}$$

此处 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的特征值。

【证明】 取 A 的属于 λ_1 的一个特征向量 α_1，则

$$A\alpha_1 = \lambda_1\alpha_1,$$

由于 $\alpha_1 \neq 0$，故

$$A\frac{\alpha_1}{\|\alpha_1\|} = \lambda_1\frac{\alpha_1}{\|\alpha_1\|}. \tag{3}$$

对列向量 $\frac{\alpha_1}{\|\alpha_1\|}$ 与 $e_1=(1, 0, \cdots, 0)'$，由定理 1.1，可找到实镜象阵 H_1，使

$$H_1\frac{\alpha_1}{\|\alpha_1\|} = e_1. \tag{4}$$

把(3)式两边左乘 $H_1(=H_1')$，并改写成：

$$(H_1'AH_1)\left(H_1\frac{\alpha_1}{\|\alpha_1\|}\right) = \lambda_1\left(H_1\frac{\alpha_1}{\|\alpha_1\|}\right).$$

由(4)式即得

$$(H_1'AH_1)e_1 = \lambda_1 e_1,$$

故

$$H_1'AH_1 = \begin{pmatrix}\lambda_1 & * \\ 0 & A_1\end{pmatrix}. \tag{5}$$

由于相似矩阵有相同的特征值，故由(5)式可知，A_1 的特征值是 $\lambda_2, \lambda_3, \cdots, \lambda_n$，再对 $n-1$ 阶阵 A_1 用归纳法，存在 $n-1$ 阶正交阵 Q_1。它是有限个 $n-1$ 阶实镜象阵的乘积，使

$$Q_1'AQ_1 = \begin{pmatrix}\lambda_2 & & & \\ & \lambda_3 & & * \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix}. \tag{6}$$

记 $Q = H_1\begin{pmatrix}1 & 0 \\ 0 & Q_1\end{pmatrix}$，则由(5)式与(6)式即得(2)式。　　证毕。

由定理 2.1 显然可得下列重要推论：

定理 2.2 对任意 n 阶实对称阵 A，必可找到正交阵 Q，使得

$$Q'AQ = [\lambda_1, \lambda_2, \cdots, \lambda_n], \tag{7}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$（显然）是 A 的特征值。

【证明】 由于A是实对称阵，故A的特征值全是实数，由定理2.1，存在正交阵Q，使成立(2)式，于是由A的对称性即得

$$Q'AQ=(Q'A'Q)'=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ * & & & \lambda_n\end{pmatrix}.$$

比较上式与(2)式的右端，即知矩阵的 * 号处的元素全是零，故得(7)式。 证毕。

由于定理2.1的证明是构造性的，故定理2.2也是构造性的，只要A的特征值能够求出。

用正交阵Q把实对称阵正交相似于对角阵，无论在理论上还是在应用上都十分重要，它的几何意义将在下一章中进一步说明。

例1 找正交阵Q，使$Q'AQ$为对角阵，此处

$$A=\begin{pmatrix}3 & -2 & -4\\ -2 & 6 & -2\\ -4 & -2 & 3\end{pmatrix}.$$

【解】 先求出A的三个特征值，由特征多项式的降阶定理，

$$\begin{aligned}|\lambda I_3-A|&=\left|(\lambda-7)I_3+\begin{pmatrix}4 & 2 & 4\\ 2 & 1 & 2\\ 4 & 2 & 4\end{pmatrix}\right|\\&=\left|(\lambda-7)I_3+\begin{pmatrix}2\\1\\2\end{pmatrix}(2,\ 1,\ 2)\right|\\&=(\lambda-7)^2(\lambda+2),\end{aligned}$$

故A的特征值为$-2,\ 7,\ 7$。

对$\lambda_1=-2$，解特征方程：$(-2I_3-A)x=0$，求得A的属于$\lambda_1=-2$的一个特征向量$\alpha_1=(2,\ 1,\ 2)'$，作单位向量

$$\beta_1=\frac{\alpha_1}{\|\alpha_1\|}=\left(\frac{2}{3},\ \frac{1}{3},\ \frac{2}{3}\right)',$$

对 β_1 与 $e_1=(1,0,0)'$，作单位(法)向量：

$$u=\frac{\beta_1-e_1}{\|\beta_1-e_1\|}=\frac{\sqrt{3}}{\sqrt{2}}\left(-\frac{1}{3},\ \frac{1}{3},\ \frac{2}{3}\right)'$$

故找到了实镜象阵：

$$H_1=I_3-2uu'=\begin{pmatrix}\frac{2}{3} & \frac{1}{3} & \frac{2}{3}\\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3}\\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3}\end{pmatrix}.$$

并且容易算得

$$H_1'AH_1=H_1AH_1=[-2,7,7],$$

也就是只经过一次找 H_1 的过程，就把 A 正交相似于对角阵，且 H_1 还是实对称阵。

注意：这种正交阵不是唯一的，例如取

$$Q_1=\begin{pmatrix}\frac{2}{3} & \frac{\sqrt{3}}{2} & \frac{\sqrt{2}}{6}\\ \frac{1}{3} & 0 & \frac{-2\sqrt{2}}{3}\\ \frac{2}{3} & \frac{-\sqrt{3}}{2} & \frac{\sqrt{2}}{6}\end{pmatrix},$$

则容易验证：$Q_1'AQ_1=[-2,7,7]$。显然 Q_1 是正交阵，但不是实对称阵，且 Q_1 不如 H_1 简洁(关于 Q_1 的求法，在第十一章中将有一般性的介绍)。

例 2 设

$$A=\begin{pmatrix}0 & 1 & 1 & -1\\ 1 & 0 & -1 & 1\\ 1 & -1 & 0 & 1\\ -1 & 1 & 1 & 0\end{pmatrix},$$

求以有理数为元素的对称、正交阵，使 $Q'AQ$ 为对角阵。

【解】 由特征多项式的降阶定理，

$$|\lambda I_4 - A| = |(\lambda-1)I_4 + \alpha\alpha'| = (\lambda-1)^3(\lambda+3),$$

其中 $\alpha=(1, -1, -1, 1)'$。

由于A与$[1, 1, 1, -3]$都是有理数域上的矩阵，故由定理2.2以及第六章§2例1可知，它们必在有理数域上相似，但按题意要求：它们在有理数域上正交相似，故每个镜象阵的元素最好是有理数，才有可能达到此目的，也就是每个镜象阵的单位(法)向量最好选取以有理数为分量的"有理列向量"。今对 $\lambda_1=1$，解特征方程$(I_4-A)x=0$，可选出一个"有理列向量"解：$\alpha_1=(1, 1, 1, 1)'$。且单位向量：

$$\beta_1=\frac{\alpha_1}{\|\alpha_1\|}=\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)'$$

也是"有理列向量"，对 β_1 与 $e_1=(1, 0, 0, 0)'$，作单位(法)向量：

$$u=\frac{\beta_1-e_1}{\|\beta_1-e_1\|}=\left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right)',$$

它仍是"有理列向量"，故得到了有理数域上的镜象阵：

$$H_1=I_4-2uu'=\frac{1}{2}\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix},$$

且容易算得：$H_1AH_1=[1, 1, 1, -3]$。

本例也是找一次实镜象阵就得到了所要的结果。另外，容易验证：正交阵

$$Q_1=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{2}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{2}\\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2}\\ 0 & 0 & \frac{3}{\sqrt{12}} & \frac{1}{2}\end{pmatrix}$$

也使 $Q_1AQ_1=[1, 1, 1, -3]$，但找 Q_1 却麻烦得多，且 Q_1 不符合本例的要求。

上面的两个例子表明，用镜象阵的方法处理这些问题，在某些时候可以收到事半功倍的效果。但须注意：不是任何实对称阵都可以只经过一次镜象阵就化为对角阵。

定理 2.2 的逆定理显然也正确，即有：如果实方阵 A 正交相似于实对角阵，则 A 必是实对称阵。这正、反两个结论说明，除了实对称阵外，再也没有其他实方阵可以正交相似于实对角阵了。

由定理 2.2 及第五章定理 2.1 显然可得：

推论 1 n 阶实对称阵有完全的特征向量系 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，且 $\alpha_i'\alpha_j=\delta_{ij}$，$i, j=1, 2, \cdots, n$。

推论 2 设 $\lambda=a+b\sqrt{-1}$ 是 n 阶实方阵 A 的任一特征值，a 与 b 分别是 λ 的实部与虚部，如果 $A+A'$ 的 n 个特征值是

$$\mu_1, \mu_2, \cdots, \mu_n,$$

则必有：

$$\frac{1}{2}\min_i \mu_i \leqslant a \leqslant \frac{1}{2}\max_i \mu_i. \tag{8}$$

【证明】 因为 $A+A'$ 是实对称阵，故由定理 2.2，存在正交阵 Q，使得

$$Q'(A+A')Q=[\mu_1, \mu_2, \cdots, \mu_n]. \tag{9}$$

设 x 是 A 的属于 λ 的特征向量，则

$$Ax=\lambda x, \tag{10}$$

对(10)式两边转置、共轭，得到

$$\bar{x}'A'=\bar{\lambda}\bar{x}'. \tag{11}$$

由(10)式与(11)式即得

$$\bar{x}'Ax=\lambda\bar{x}'x,$$
$$\bar{x}'A'x=\bar{\lambda}\bar{x}'x,$$

于是

$$\bar{x}'(A+A')x=(\lambda+\bar{\lambda})\bar{x}'x=2a\bar{x}'x. \tag{12}$$

令 $x=Qy$，其中 $y=(y_1, y_2, \cdots, y_n)'$ 是复向量，由(9)式，(12)式可化为

$$\bar{y}'[\mu_1, \mu_2, \cdots, \mu_n]y = 2a\bar{y}'y,$$

也即

$$\sum_{i=1}^{n} \mu_i \bar{y}_i y_i = 2a \sum_{i=1}^{n} \bar{y}_i y_i,$$

由上式可知

$$(\min_i \mu_i) \sum_{i=1}^{n} |y_i|^2 \leqslant 2a \sum_{i=1}^{n} |y_i|^2 \leqslant (\max_i \mu_i) \sum_{i=1}^{n} |y_i|^2. \qquad (13)$$

因为 $x \neq 0$, Q 是非异阵，故 $y \neq 0$，于是 $\sum_{i=1}^{n} |y_i|^2 \neq 0$，所以由(13)式即得(8)式。 证毕。

* **例 3** （**Cantoni-Butler**，1976）设 $A = (a_{ij})_{n\times n}$，$J_n = (e_n, e_{n-1}, \cdots, e_1)$，$e_j$ 是标准单位向量，如果 $A = J_n A J_n$，即

$$a_{ij} = a_{n+1-i, n+1-j}, \ i, j = 1, 2, \cdots, n,$$

则称 A 为中心对称阵，如果进一步，A 还是实对称阵，则称 A 为实对称的中心对称阵，并简称 A 为 SC 阵①。

当 $n = 2k$ 时，由定义可知，SC 阵可写成

$$A = \begin{pmatrix} B & J_k C J_k \\ C & J_k B J_k \end{pmatrix}, \qquad (14)$$

并且 B 是实对称阵，$C' = J_k C J_k$。

求证：$2k$ 阶 SC 阵 A 有如下性质：

(i) $B + J_k C$, $B - J_k C$ 有完全的特征向量系；

(ii) A 的完全特征向量系为如下 $2k$ 个线性无关向量：

$$\begin{pmatrix} \alpha_i \\ -J_k \alpha_i \end{pmatrix}, \begin{pmatrix} \beta_i \\ J_k \beta_i \end{pmatrix}, i = 1, 2, \cdots, k,$$

而 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 为 $B - J_k C$ 的某个完全特征向量系；$\beta_1, \beta_2, \cdots, \beta_k$ 为 $B + J_k C$ 的某个完全特征向量系。

【证明】 易证 (i)，因为由 $C' = J_k C J_k$ 可知，$C' J_k = J_k C$，故 $J_k C$ 是实对称阵，因而 $B \pm J_k C$ 都是实对称阵，由定理 2.2 的推论 1，它们有完全的特征向量系。

---

① SC 阵在应用数学中有较多的应用，诸如信息论、线性系统论、线性估计论、数值分析等学科会用到它。

再证(ii)，作 $2k$ 阶正交阵

$$Q=\frac{1}{\sqrt{2}}\begin{pmatrix} I_k & I_k \\ -J_k & J_k \end{pmatrix},$$

则容易算得

$$Q'AQ=[B-J_kC,\ B+J_kC] \tag{15}$$

(Q 可用第一章中同时求 P 与 Q 的方法找出，这里 $Q'=P$)。又由(i)，$B\pm J_kC$ 是实对称阵，故由定理 2.2，存在正交阵 U 与 V，使

$$U'(B-J_kC)U=[\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_k], \tag{16}$$

$$V'(B+J_kC)V=[\lambda_{k+1},\ \lambda_{k+2},\ \cdots,\ \lambda_{2k}], \tag{17}$$

此处 $\lambda_1,\ \lambda_2,\ \cdots,\lambda_k$ 与 $\lambda_{k+1},\ \lambda_{k+2},\cdots,\lambda_{2k}$ 分别是 $B-J_kC$ 与 $B+J_kC$ 的特征值。于是由(15)、(16)、(17)式可得

$$\begin{pmatrix} U & O \\ O & V \end{pmatrix}' Q'AQ \begin{pmatrix} U & O \\ O & V \end{pmatrix}=[\lambda_1,\ \cdots,\ \lambda_k,\ \lambda_{k+1},\ \cdots,\ \lambda_n]. \tag{18}$$

由(18)式可知 $Q\begin{pmatrix} U & O \\ O & V \end{pmatrix}$ 的 $2k$ 个列，或者，

$$\sqrt{2}\,Q\begin{pmatrix} U & O \\ O & V \end{pmatrix}=\begin{pmatrix} U & V \\ -J_kU & J_kV \end{pmatrix},$$

的 $2k$ 个列是 A 的 $2k$ 个特征向量，它们组成 A 的完全特征向量系，即

$$\begin{pmatrix} \alpha_i \\ -J_k\alpha_i \end{pmatrix}、\begin{pmatrix} \beta_i \\ J_k\beta_i \end{pmatrix}, \qquad i=1,\ 2,\ \cdots,\ k$$

是 A 的完全特征向量系，此处

$$U=(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_k),\ V=(\beta_1,\ \beta_2\cdots,\ \beta_k),$$

而 $\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_k;\beta_1,\ \beta_2,\ \cdots,\ \beta_k$ 分别是 $B-J_kC$ 与 $B+J_kC$ 的完全特征向量系。 证毕。

二、任意实方阵正交相似的矩阵

上一段已证明，除了实对称阵外，再也没有其他实方阵可以正交相似于对角阵，那么，任意实方阵是否可正交相似于一个比对角阵复杂一点，但比原矩阵要简单一点的实方阵呢？本段将讨论这一

问题，先证：

引理 1 设 $2s$ 阶实方阵 A 的特征值为

$$\lambda_i = a_i \pm \sqrt{-1} b_i,$$

a_i 与 b_i 是 λ_i 的实部与虚部，且 $b_i \neq 0$, $i=1, 2, \cdots, s$（即 A 没有实特征值），则必存在正交阵 Q，使

$$Q'AQ = \begin{pmatrix} A_1 & & * & \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_s \end{pmatrix}, \tag{19}$$

其中 A_i 都是 2 阶阵，它的特征值是

$$a_i \pm b_i\sqrt{-1}, \qquad i=1, 2, \cdots, s.$$

【证明】 对 s 用归纳法。$s=1$ 时引理 1 显然正确，设引理 1 对 $s-1$ 成立，今证引理 1 对 s 也成立。令 $\alpha_1 + \sqrt{-1}\beta_1$ 是 A 的属于 $\lambda_1 = a_1 + \sqrt{-1}\, b_1$ 的特征向量，α_1、β_1 都是 n 维实的列向量，则由

$$A(\alpha_1 + \sqrt{-1}\,\beta_1) = (a_1 + \sqrt{-1}\, b_1)(\alpha_1 + \sqrt{-1}\,\beta_1),$$

可得

$$A\alpha_1 = a_1\alpha_1 - b_1\beta_1, \tag{20}$$

$$A\beta_1 = b_1\alpha_1 + a_1\beta_1, \tag{21}$$

由(20)与(21)可知，α_1, β_1 线性无关。因如不然，则

$$\beta_1 = k_1\alpha_1, \tag{22}$$

且 k_1 是非零实数（否则 $\beta_1 = 0$，再由(21)式，$b_1\alpha_1 = 0$，但 $b_1 \neq 0$ 故 $\alpha_1 = 0$。此为矛盾），把(20)式两边乘以 k_1，并以(22)式代入，即得

$$k_1A\alpha_1 = k_1a_1\alpha_1 - k_1^2b_1\alpha_1. \tag{23}$$

又由(21)与(22)式可得

$$k_1A\alpha_1 = A\beta_1 = k_1a_1\alpha_1 + b_1\alpha_1. \tag{24}$$

由(23)与(24)式可知

$$(1+k_1^2)b_1\alpha_1 = 0. \tag{25}$$

但 $1+k_1^2 \neq 0$, $\alpha_1 \neq 0$（否则由(22)式，$\beta_1 = 0$。又会产生矛盾），故由(25)式可得 $b_1 = 0$，这与假设相矛盾，所以 α_1, β_1 必线性无关。

作 $2s$ 阶非异阵：$T = (\alpha_1, \beta_1, \delta_3, \cdots, \delta_{2s})$，由第三章定理 4.1

可知这种 T 必然存在，于是

$$
\begin{aligned}
AT &= (A\alpha_1,\ A\beta_1,\ A\delta_3,\ \cdots,\ A\delta_{2s}) \\
&= (a_1\alpha_1 - b_1\beta_1,\ b_1\alpha_1 + a_1\beta_1,\ A\delta_3,\ \cdots,\ A\delta_{2s}) \\
&= (T(a_1e_1 - b_1e_2),\ T(b_1e_1 + a_1e_2),\ A\delta_3,\ \cdots,\ A\delta_{2s}) \\
&\quad (\text{因 } \alpha_1 = Te_1,\ \beta_1 = Te_2) \\
&= T(a_1e_1 - b_1e_2,\ b_1e_1 + a_1e_2,\ T^{-1}A\delta_3,\ \cdots,\ T^{-1}A\delta_{2s}) \\
&= T\begin{pmatrix} \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix} & * \\ 0 & * \end{pmatrix}. \qquad (26)
\end{aligned}
$$

记 $B_1 = \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}$，易知 B_1 的特征值恰好是 $a_1 \pm \sqrt{-1}\, b_1$。 把(26)式写成：

$$
AT = T\begin{pmatrix} B_1 & * \\ 0 & * \end{pmatrix}, \qquad (27)
$$

作 T 的 Q_1R 分解：$T = Q_1R$，其中 Q_1 是正交阵，R 是非异上三角阵，代入(9)式即得

$$
AQ_1R = Q_1R\begin{pmatrix} B_1 & * \\ 0 & * \end{pmatrix},
$$

也即

$$
\begin{aligned}
Q_1'AQ_1 &= R\begin{pmatrix} B_1 & * \\ 0 & * \end{pmatrix}R^{-1} \\
&= \begin{pmatrix} R_1 & * \\ 0 & * \end{pmatrix}\begin{pmatrix} B_1 & * \\ 0 & * \end{pmatrix}\begin{pmatrix} R_1^{-1} & * \\ 0 & * \end{pmatrix} \\
&= \begin{pmatrix} R_1B_1R_1^{-1} & * \\ 0 & * \end{pmatrix} = \begin{pmatrix} A_1 & * \\ 0 & B \end{pmatrix}, \qquad (28)
\end{aligned}
$$

其中 $A_1 = R_1BR_1^{-1}$ 是 2 阶实方阵，它与 B_1 有相同的特征值，故 A_1 的特征值仍是 $a_1 \pm \sqrt{-1}\, b_1$。 由(28)式得到

$$
|\lambda I_{2s} - A| = |\lambda I_2 - A| \cdot |\lambda I_{2s-2} - B|,
$$

所以 B 的 $2s-2$ 个特征值恰好是：

$$
a_2 \pm \sqrt{-1}\, b_2,\ \cdots,\ a_s \pm \sqrt{-1}\, b_s.
$$

由归纳法假设，存在 $2s-2$ 阶正交阵 Q_2，使

$$Q_2'BQ_2=\begin{pmatrix} A_2 & & & \\ & A_3 & & * \\ & & \ddots & \\ 0 & & & A_s \end{pmatrix}, \tag{29}$$

而 A_i 的特征值是 $a_i\pm\sqrt{-1}b_i$, $i=2,\cdots,s$。

记 $Q=[I_2, Q_2]$，则 Q 也是正交阵，且由(28)式与(29)式即可得出(1)式。　　　　证毕。

定理 2.3　设 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 是 n 阶实方阵 A 的实特征值，$a_i\pm\sqrt{-1}\,b_i$ 是 A 的复特征值（a_i, b_i 是实数，$b_i\neq0$），$i=1, 2, \cdots, s$; $r+2s=n$，则必存在正交阵 Q，使

$$Q'AQ=\begin{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ 0 & & & \lambda_r \end{pmatrix} & * \\ 0 & \begin{pmatrix} A_1 & & & \\ & A_2 & & * \\ & & \ddots & \\ 0 & & & A_s \end{pmatrix} \end{pmatrix}, \tag{30}$$

其中 A_i 都是 2 阶阵，它们的特征值是

$$a_i\pm\sqrt{-1}\,b_i,\ i=1, 2, \cdots, s.$$

【证明】　因为 A 是实方阵，$\lambda_1, \lambda_2, \cdots, \lambda_r$ 是实数，故与定理 2.1 的证法相同，可证存在有限个实镜象阵 H_1、H_2、$\cdots$、H_l，$l\leqslant r$，使

$$H_l'\cdots H_1'AH_1\cdots H_l=\begin{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ 0 & & & \lambda_r \end{pmatrix} & * \\ 0 & B \end{pmatrix}, \tag{31}$$

易知 B 的特征值恰好是 $a_i\pm\sqrt{-1}\,b_i$, $i=1, 2, \cdots, s$. 故由引理 1，存在 $2s$ 阶正交阵 P，使

$$P'BP=\begin{pmatrix} A_1 & & & \\ & A_2 & & * \\ & & \ddots & \\ 0 & & & A_s \end{pmatrix}. \tag{32}$$

记 $Q=H_1H_2\cdots H_t\cdot[I_r,\ P]$，则 Q 仍是正交阵，且由 (31) 式与 (32)式即得(30)式。 证毕。

*三、实正规阵的正交相似标准形

理论上与应用上常需考虑一种被称为实正规阵的正交相似标准形，所谓**实正规阵** A 指的是满足 $AA'=A'A$ 的实方阵。例如，实对称阵、反对称(实方)阵、正交阵、拟正交阵 ($AA'=bI$, b 是正实数)等等都是实正规阵的例子。

引理 2 设 A、C 是方阵(未必同阶)，且 $\begin{pmatrix} A & B \\ O & C \end{pmatrix}$ 是实正规阵，则必 $B=0$，且 A、C 仍是实正规阵。

【证明】 由假设，

$$\begin{pmatrix} A' & O \\ B' & C' \end{pmatrix}\begin{pmatrix} A & B \\ O & C \end{pmatrix}=\begin{pmatrix} A & B \\ O & C \end{pmatrix}\begin{pmatrix} A' & O \\ B' & C' \end{pmatrix}, \tag{33}$$

比较(33)式两边分块阵乘积的元素，可得

$$A'A=AA'+BB', \tag{34}$$

$$B'B+C'C=CC', \tag{35}$$

对(35)的两边的方阵取它们的迹，则得

$$\operatorname{tr}(B'B)+\operatorname{tr}(C'C)=\operatorname{tr}(CC'),$$

但是 $\operatorname{tr}(CC')=\operatorname{tr}(C'C)$，故上式化为 $\operatorname{tr}(B'B)=0$，因为 B 是实方阵，故 $B=0$，把它代入(34)、(35)式，即知 A、C 也是实正规阵。

证毕。

定理 2.4 设 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 是 n 阶实正规阵 A 的实特征值，$a_i\pm\sqrt{-1}\,b_i$ 是 A 的复特征值，a_i 与 b_i 都是实数，且

$$b_i\neq 0,\ i=1,\ 2,\ \cdots,\ s;\quad r+2s=n,$$

则必存在正交阵 Q，使

$$Q'AQ=[\lambda_1, \lambda_2, \cdots, \lambda_r, A_1, A_2, \cdots, A_s], \tag{36}$$

而

$$A_i=\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix},\ i=1,\ 2,\ \cdots,\ s. \tag{37}$$

【证明】 由定理 2.3，对任何实方阵 A，必可找到正交阵 Q，使成立(30)式。今 A 是实正规阵，故易知 $Q'AQ$ 也是正规阵，由引理 2，(30)式右边的方阵的右上角 * 号处应是零，且(30)式的主对角块(元)都是实正规阵，故由归纳法并多次运用引理 2，可知(30)式右边方阵的所有 * 号处的元素全是零，且 A_1、A_2、…、A_s 分别是特征值为 $a_i \pm \sqrt{-1}\, b_i$ 的 2 阶实正规阵，$i=1, 2, \cdots, s$。而 2 阶实正规阵 A_i 必是下面的形状：

$$A_i = \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}, \quad i=1, 2, \cdots, s$$

(这由简单的矩阵乘法运算便知，留给读者作为练习)。所以对实正规阵 A 来说，(30)式就化为(36)式的形状，并且 A_i 必是(37)的形状。 证毕。

称(36)式(包括(37)式)为实正规阵的**正交相似标准形**。

定理 2.4 的逆显然也正确：即正交相似于(36)式右边方阵(A_i 是形如(37)的方阵)的实方阵必是实正规阵。这个结论与定理 2.4 合起来说明：除了实正规阵外，再也没有其他实方阵可以正交相似于形如(36)式右边的分块对角阵了。

推论 1 相似的两个实正规阵必正交相似。

【证明】 设 A、B 是两个实正规阵，由假设它们是相似的，故它们有相同的特征值，设 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 是它们的所有实特征值；$a_i \pm \sqrt{-1}\, b_i$，$i=1, 2, \cdots, s$ 是它们的所有的复特征值，则由定理 2.4，必存在正交阵 Q_1 与 Q_2，使

$$Q_1'AQ_1 = \left[\lambda_1, \lambda_2, \cdots, \lambda_r, \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}\right],$$

$$Q_2'BQ_2 = \left[\lambda_1, \lambda_2, \cdots, \lambda_r, \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}\right],$$

故由上两式即得

$$B = Q_2Q_1'AQ_1Q_2' = (Q_1Q_2')'A(Q_1Q_2')。$$

但 Q_1Q_2' 仍是正交阵，所以 A 与 B 正交相似。 证毕。

推论 2 任何实正规阵必可分解为两个实对称阵的乘积，且其中有一个是正交、对称阵(即实对合阵)。

【证明】 设A是$r+2s$阶实正规阵，$\lambda_1, \lambda_2, \cdots, \lambda_r$是$A$的所有实特征值，$a_i \pm \sqrt{-1}\, b_i$是$A$的所有复特征值，$i=1, 2, \cdots, s$，则

$$A = Q\left[\lambda_1, \lambda_2, \cdots, \lambda_r, \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}\right]Q', \tag{38}$$

其中Q是正交阵，因为

$$\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix} = \begin{pmatrix} b_i & a_i \\ a_i & -b_i \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = B_i J_2, \quad i=1, 2, \cdots, s, \tag{39}$$

此处

$$B_i = \begin{pmatrix} b_i & a_i \\ a_i & -b_i \end{pmatrix} \qquad J_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

分别是实对称阵与对称、正交阵，故由(38)式与(39)式即得

$$\begin{aligned} A &= (Q[\lambda_1, \cdots, \lambda_r, B_1, \cdots, B_s]Q')(Q[I_r, J_2, \cdots, J_2]Q') \\ &= SP. \end{aligned}$$

易知$S = Q[\lambda_1, \cdots, \lambda_r, B_1, \cdots, B_s]Q'$是对称阵，而

$$P = Q[I_r, J_2, \cdots, J_2]Q'$$

是对称、正交阵. 证毕.

习　　题

1. 对下列对称阵A，找正交阵Q，使$Q'AQ$为对角阵：

(i) $A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$;　　(ii) $A = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$;

(iii) $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$;　　(iv) $A = \begin{pmatrix} 2 & -4 & 1 \\ -4 & 17 & -4 \\ 1 & -4 & 2 \end{pmatrix}$;

(v) $A=\begin{pmatrix}\sqrt{2} & -1 & 0\\ -1 & 1 & -\sqrt{2}\\ 0 & -\sqrt{2} & 0\end{pmatrix}$

(vi) $A=\begin{pmatrix}1 & 1 & 0 & -1\\ 1 & 1 & -1 & 0\\ 0 & -1 & 1 & 1\\ -1 & 0 & 1 & 1\end{pmatrix}$;

(vii) $A=\begin{pmatrix}-1 & -3 & 3 & -3\\ -3 & -1 & -3 & 3\\ 3 & -3 & -1 & -3\\ -3 & 3 & -3 & -1\end{pmatrix}$;

(viii) $A=\begin{pmatrix}1 & -1 & 0 & 0\\ -1 & 1 & 0 & 0\\ 0 & 0 & 3 & -1\\ 0 & 0 & -1 & 2\end{pmatrix}$.

2. 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 n 阶实对称阵 A 的特征值，求证：

(i) $A=\lambda_1E_1+\lambda_2E_2+\cdots+\lambda_nE_n$.

而 $E_iE_j=\delta_{ij}E_j$, $i, j=1, 2\cdots, n$.

(ii) $A=\lambda_1\alpha_1\alpha_1'+\lambda_2\alpha_2\alpha_2'+\cdots+\lambda_n\alpha_n\alpha_n'$

而 $\alpha_i'\alpha_j=\delta_{ij}$, $i, j=1, 2, \cdots, n$.

3. 求证：秩为 r 的实对称阵必可分解为 r 个秩为1的实对称阵之和.

4. 设 A 与 B 是 n 阶实对称阵，且 $AB=BA$，求证：必存在正交阵 Q，使 $Q'AQ=[\lambda_1, \lambda_2, \cdots, \lambda_n]$，$Q'BQ=[\mu_1, \mu_2, \cdots, \mu_n]$. 此处 λ_i 与 μ_i 分别是 A 与 B 的特征值，$i=1, 2, \cdots, n$.

(提示：设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 全不相同，且 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 中有 n_1 个 λ_1，n_2 个 λ_2，$\cdots$，n_s 个 λ_s，则存在正交阵 P，使 $P'AP=[\lambda_1I_{n_1}, \lambda_2I_{n_2}, \cdots, \lambda_sI_{n_s}]$，然后用假设条件：$AB=BA$，并对 $P'BP$ 再应用定理2.2.)

§ 3 复方阵的酉相似、Schur 定理及其应用

一、Schur 定理

定义 设 A 与 B 都是复方阵，如果存在酉阵 Q，使

$$Q^{-1}AQ=B \ (\text{即 } \overline{Q}'AQ=B), \tag{1}$$

则称 A 与 B **酉相似**，或称 A 酉相似于 B。

定理 3.1 (Schur) 任何 n 阶复方阵必酉相似于上三角阵，即存在酉阵 Q，使得

$$\overline{Q}'AQ=\begin{pmatrix}\lambda_1 & c_{12} & c_{13} & \cdots & c_{1n}\\ & \lambda_2 & c_{23} & \cdots & c_{2n}\\ & & \lambda_3 & \cdots & c_{3n}\\ & & & \ddots & \vdots\\ & & & & \lambda_n\end{pmatrix}, \tag{2}$$

其中 $\lambda_1, \lambda_2, \cdots\cdots, \lambda_n$ (显然)是 A 的(复)特征值。

【证明】 因为对任何复方阵 A，必存在非异复方阵 T，使得

$$T^{-1}AT=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix}, \tag{3}$$

(第五章定理 2.2)。由于 T 是非异阵，故由定理 1.4，

$$T=QR, \tag{4}$$

其中 Q 是酉阵，R 是非异上三角阵，把(4)式代入(3)式，并经过整理后即得

$$Q^{-1}AQ=R\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix}R^{-1}, \tag{5}$$

由于 R 是上三角阵，故 R^{-1} 也是上三角阵，故(5)式右边仍是一个上三角阵，故(5)式可写成：

$$Q^{-1}AQ=\begin{pmatrix}b_{11} & & & \\ & b_{22} & & * \\ & & \ddots & \\ 0 & & & b_{nn}\end{pmatrix}, \tag{6}$$

由于相似阵有相同的特征值，故 (6) 式的 $b_{11}, b_{22}, \cdots, b_{nn}$ 就是 A

的特征值，故得(2)式。　　证毕。

推论 1　Hermite 阵(或实对称阵)A 的特征值全是实数。

【证明】　这个已证明过的结论由 Schur 定理几乎立刻可以看出：因为对任意 A，存在酉阵 Q，使成立(2)式，即

$$\overline{Q}'AQ=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & * \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix},\tag{7}$$

对上式两边转置、取共轭，即得

$$\overline{Q}'\overline{A}'Q=\begin{pmatrix}\bar{\lambda}_1 & & & \\ & \bar{\lambda}_2 & & 0 \\ & ** & \ddots & \\ & & & \bar{\lambda}_n\end{pmatrix}.\tag{8}$$

由假设 $\overline{A}'=A$，并比较(7)式与(8)式的右边，即得 $\lambda_i=\bar{\lambda}_i$，故 λ_i 是实数，$i=1, 2, \cdots, n$。　　证毕。

从(7)式与(8)式右边的矩阵还可看出，当 $A=\bar{A}'$ 时，* 号处与** 号处的元素全为零，故还可得：

推论 2　对 Hermite 阵 A，必存在酉阵 Q，使 $Q'AQ$ 为实对角阵。

推论 2 之逆也正确；若复方阵 A 酉相似于实对角阵，则 A 必是 Hermite 阵。由此及推论 2 可知，除了 Hermite 阵外，再也没有其他复方阵能酉相似于实对角阵了。

除了 Hermite 阵可酉相似于实对角阵外，还有何种复方阵可酉相似于复对角阵？这一问题将在选做题第 13 题中给出答案。Schur 定理另外一个十分有用的推论是下面的：

定理 3.2　设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是任意 n 阶阵 $A=(a_{ij})_{n\times n}$ 的特征值，则

$$\sum_{i=1}^{n}|\lambda_i|^2\leqslant\sum_{i,j=1}^{n}|a_{ij}|^2.\tag{9}$$

【证明】　由 Schur 定理，对任何 n 阶复方阵，必存在酉阵 Q，

使成立(2)式，在(2)式两边取共轭、转置，即得

$$\overline{Q}'\overline{A}'Q=\begin{pmatrix}\bar{\lambda}_1 & & & & \\ \bar{c}_{12} & \bar{\lambda}_2 & & 0 & \\ \bar{c}_{13} & \bar{c}_{23} & \bar{\lambda}_3 & & \\ \cdots & \cdots & \cdots & \ddots & \\ \bar{c}_{1n} & \bar{c}_{2n} & \bar{c}_{3n} & \cdots & \bar{\lambda}_n\end{pmatrix}. \tag{10}$$

由(2)式与(10)式可得关系式：

$$\overline{Q}'A\overline{A}'Q=\begin{pmatrix}|\lambda_1|^2+\sum_{j=2}^{n}|c_{1j}|^2 & & & * \\ & |\lambda_2|^2+\sum_{j=3}^{n}|c_{2j}|^2 & & \\ * & & \ddots & \\ & & & |\lambda_n|^2\end{pmatrix}, \tag{11}$$

$$\overline{Q}'\overline{A}'AQ=$$

$$\begin{pmatrix}|\lambda_1|^2 & & & & \\ & |\lambda_2|^2+|c_{12}|^2 & & & \\ & & |\lambda_3|^2+\sum_{i=1}^{2}|c_{i3}|^2 & & * \\ & & & \ddots & \\ * & & & & |\lambda_n|^2+\sum_{i=1}^{n-1}|c_{in}|^2\end{pmatrix}. \tag{12}$$

(11)式与(12)式右边的方阵中 * 号处的元素(非主对角元)不必写出，因为用不到它们，对(12)式两边的方阵取迹，则得

$$\sum_{i=1}^{n}|\lambda_i|^2+\sum_{1\le i<j\le n}|c_{ij}|^2=\operatorname{tr}(\overline{A}'A)=\sum_{i,j=1}^{n}|a_{ij}|^2 \tag{13}$$

(称(13)式为 **Schur 恒等式**)，由(13)式显然可得(9)式，

证毕。

称(9)式为 **Schur 不等式**。Schur 不等式是线性代数中一个极端重要的不等式，它在行列式理论、矩阵的秩以及特征值的估计

及分布理论中都有它的应用（见下一段以及本章选做题第 8 题到第 12 题）。

*二、方阵的非异性的判定

设 $A=(a_{ij})_{n\times n}$ 是复方阵，在第三章中已证明，当 A 是严格对角占优阵，即 $|a_{ii}|>\sum\limits_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|$，$i=1, 2, \cdots, n$ 时，A 是非异阵（Lévy–Hadamard 定理），从那时起直到本世纪七十年代，陆续提供了不少有关判定 A 是非异阵的充分条件，它们中的大多数是把 Lévy–Hadamard 的 n 个条件减弱，或作个别的更换，也都是在一个个主对角元占优的条件下进行，本段将从主对角元之和占优（即方阵的迹占优）这一总体考虑问题，应用 Schur 定理，获得判断方阵非异性的一个充分条件，这就是下面的：

定理 3.3 设 $A=(a_{ij})_{n\times n}$ 是复方阵，如果

$$\left|\sum_{i=1}^{n}a_{ii}\right|^2>(n-1)\sum_{i,j}|a_{ij}|^2, \tag{14}$$

则 A 必是非异阵。

定理 3.3 是下面的定理的一个直接推论，这就是

定理 3.4 设 $r(A)$ 是任意 n 阶阵 $A=(a_{ij})_{n\times n}$ 的秩，则

$$r(A)\geqslant\frac{|\mathrm{tr}A|^2}{\mathrm{tr}(A\bar{A}')}=\frac{\left|\sum\limits_{i=1}^{n}a_{ii}\right|^2}{\sum\limits_{i,j}^{n}|a_{ij}|^2}. \tag{15}$$

【证明】 记 $r(A)=r$。设 A 的特征值是 $\lambda_1, \lambda_2, \cdots, \lambda_n$，由 Schur 定理，存在酉阵 Q，使

$$\bar{Q}'AQ=\begin{pmatrix}\lambda_1 & & * \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n\end{pmatrix}. \tag{16}$$

易知 A 的非零特征值的个数 s 不会超过 r，即 $s\leqslant r$。不然的话，若 $s>r$，则(16)式右边的方阵中至少有一个 s 阶子式不等于零，从而

产生 $r=\mathrm{r}(A)=\mathrm{r}(\overline{Q}'AQ)\geqslant s$ 的矛盾。

为书写简便起见，不妨设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 不等于零，由第二章的 Cauchy-Schwarz 不等式，

$$\left|\sum_{i=1}^{n}\lambda_i\right|^2=\left|\sum_{i=1}^{s}\lambda_i\right|^2\leqslant\left(\sum_{i=1}^{s}1^2\right)\left(\sum_{i=1}^{s}|\lambda_i|^2\right)=s\sum_{i=1}^{n}|\lambda_i|^2,$$

故由 Schur 不等式，上式可化为：

$$|\mathrm{tr}A|^2\leqslant s\sum_{i,j=1}^{n}|a_{ij}|^2=s\,\mathrm{tr}(A\overline{A}')\leqslant r\,\mathrm{tr}(A\overline{A}'),$$

上式即(21)式。 证毕。

称(15)式为方阵的秩的下界估计式，它只借助于已给的方阵 A 的元素 a_{ij} 就可初步判断 A 的秩至少应是多少。

由定理 3.4 显然可得定理 3.3。今举两例：

例 1 求方阵

$$A=\begin{pmatrix}4&3&1\\1&4&2\\1&2&4\end{pmatrix}$$

的秩。

因为由(15)式可得

$$\mathrm{r}(A)\geqslant\frac{(12)^2}{68}=\frac{144}{68}>2,$$

故 A 的秩为 3。

例 2 求证：方阵

$$A=\begin{pmatrix}5&0&1&0&1&-1\\1&5&-1&-1&-1&0\\-1&-1&4&-1&-1&-1\\0&1&0&5&1&-1\\2&0&-2&-1&5&1\\0&-1&0&1&0&5\end{pmatrix}$$

是非异阵。

【证明】 容易算出，$\left|\sum_{i=1}^{n}a_{ii}\right|^2=841$，$\sum_{i,j=1}^{n}|a_{ij}|^2=168$，而

$$841>(6-1)\cdot 168,$$

故由定理 3.3 可知A是非异阵。 证毕。

例 2 的A不是严格对角占优阵（第 3 行第 3 列的元素小于这一行其他元素的绝对值之和，第 5 行亦然），但A的主对角元之和（的绝对值）这一“总体”仍“占优”，故仍可判断它为非异阵。

习　题

1. 对下列方阵A，找酉阵Q，使$\overline{Q}'AQ$为上三角阵

(i) $A=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$； (ii) $A=\begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix}$.

2. 用 Schur 定理证明：

(i) $|A|=\lambda_1\lambda_2\cdots\cdots\lambda_n$，此处$\lambda_1, \lambda_2, \cdots\cdots, \lambda_n$是$A$的特征值；

(ii) Hermite 阵的行列式是实数。

3. 用 Schur 定理证明：反对称阵、斜 Hermite 阵（即$\overline{A}'=-A$）的特征值的实部等于零。

4. 用 Schur 定理证明：正交阵、酉阵的任一特征值的绝对值（模）等于 1。

5. 找出拟酉阵的特征值所满足的条件。

6. 如果A是复方阵，且$\overline{A}'=kA$，$k\neq 1$，则称A为拟对称阵。求证：拟对称阵的特征值是零或纯虚数。

7. 设

$$A=\begin{pmatrix} 2 & 1-\sqrt{2} \\ 1+\sqrt{2} & 2 \end{pmatrix},$$

找出酉阵Q，使$\overline{Q}'AQ$是实对角阵。

8. 设A是n阶复正规阵，$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的特征值，求证：

(i) $A=\lambda_1\alpha_1\overline{\alpha}_1'+\lambda_2\alpha_2\overline{\alpha}_2'+\cdots+\lambda_n\alpha_n\overline{\alpha}_n'$，

其中α_i都是n维列向量，且

$$\overline{\alpha}_i'\alpha_j=\delta_{ij},\ i, j=1, 2, \cdots, n.$$

称上述分解式为**谱分解式**。

(ii) 必存在n维单位列向量α_j（即$\overline{\alpha}_j'\alpha_j=1$），使

$$\lambda_j=\overline{\alpha}_j'A\alpha_j,\quad 1\leqslant j\leqslant n.$$

选 做 题

1. 设 α,β 是两个长度相等的 n 维实的列向量，并且 $\alpha\neq\beta$，求证：存在唯一的一个实镜象阵 H，使 $H\alpha=\beta$。

2. 设 α,β 是两个长度相等的 n 维复的列向量，如果存在复镜象阵 H，使 $H\alpha=\beta$，求证：$\bar{\alpha}'\beta$ 必是实数。

3. 设 A 是 n 阶正交阵，求证：

(i) 必存在正交阵 Q，使

$$Q'AQ=\left[I_l,\ -I_t,\begin{pmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{pmatrix},\cdots,\begin{pmatrix}\cos\theta_s & \sin\theta_s\\ -\sin\theta_s & \cos\theta_s\end{pmatrix}\right],$$

称上式右边的分块对角阵为**正交阵的正交相似标准形**；

(ii) A 必可分解为两个实对称、正交阵(因而是实对合阵)的乘积；

(iii) A 与 A^{-1} 在实数域上相似；

(iv) A 或正交相似于 $\mathrm{adj}\,A$ 或正交相似于 $-\mathrm{adj}\,A$；

(v) $|A|=-1$ 的充要条件是：A 有奇数个特征值是 -1。

4. 设 A 是反对称(实方)阵，求证：

(i) 必存在正交阵 Q，使

$$Q'AQ=\left[0,\ \begin{pmatrix}0 & b_1\\ -b_1 & 0\end{pmatrix},\ \cdots,\ \begin{pmatrix}0 & b_s\\ -b_s & 0\end{pmatrix}\right],$$

此处 $b_i\sqrt{-1}$(显然)是 A 的特征值，称上式右边的分块对角阵为**反对称阵的正交相似标准形**；

(ii) $|A|=(f(\lambda_1,\lambda_2,\cdots,\lambda_l))^2$，

此处 $f(x_1,x_2,\cdots,x_l)$ 是实系数多元多项式；

(iii) $|A|\geqslant 0$；

(iv) 当 A 的阶是奇数时，则 $|A|=0$；

(v) 当 A 是偶数阶非异阵时，则 A^{-1} 也是反对称阵；

(vi) $I+A$ 的主子式全大于 1。

5. 求拟正交阵 $A(A'A=bI,\ b>0)$ 的正交相似标准形。

6. 记 $a=\max\limits_{1\leqslant i,j\leqslant n}|a_{ij}|$，应用下面的 Cauchy 不等式：

$$(b_1b_2\cdots b_n)^{\frac{1}{n}}\leqslant\frac{b_1+b_2+\cdots+b_n}{n}\ (b_i\geqslant 0,\ i=1,2,\cdots,n)$$

(即几何平均值不大于算术平均值)以及 Schur 不等式证明：

$$|\det A|\leqslant n^{n/2}a^n,$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是A的特征值，$A=(a_{ij})_{n\times n}$，$\det A$ 表示A的行列式。

7. 设 $\lambda_i=a_i+b_i\sqrt{-1}$ 是 n 阶阵 $A=(a_{ij})_{n\times n}$ 的特征值，a_i 与 b_i 分别是 λ_i 的实部与虚部，$i=1, 2, \cdots, n$，求证：

(i) $\sum\limits_{i=1}^{n}|a_i|^2\leqslant\sum\limits_{i,j=1}^{n}\left|\dfrac{a_{ij}+\overline{a_{ji}}}{2}\right|^2$;

(ii) $\sum\limits_{i=1}^{n}|b_i|^2\leqslant\sum\limits_{i,j=1}^{n}\left|\dfrac{a_{ij}-\overline{a_{ji}}}{2}\right|^2$.

8. 设M是 $m\times n$ 复矩阵，如果

$|\mathrm{tr}(M\overline{M}')|^2>(n-1)\mathrm{tr}((M\overline{M}')^2)$，当 $m\leqslant n$，

(或者，$|\mathrm{tr}(\overline{M}'M)|^2>(n-1)\mathrm{tr}((\overline{M}'M)^2)$，当 $m>n$，)

求证：$|M\overline{M}'|>0$ (或者 $|\overline{M}'M|>0$)。

9. 设A是 n 阶复方阵，对任一 k，$1\leqslant k\leqslant n-1$，把$A$如下分块：

$$A=\begin{pmatrix} A_k & B_k \\ C_k & D_k \end{pmatrix},$$

其中 A_k 是A的 k 阶顺序主子阵，记

$$l_k=\mathrm{tr}(A_k\overline{A}_k')+\mathrm{tr}(D_k\overline{D}_k')+2\sqrt{\mathrm{tr}(B_k\overline{B}_k')\mathrm{tr}(C_k\overline{C}_k')},$$

$$l=\min_k l_k.$$

(i) 求证：$l\leqslant\mathrm{tr}(A\overline{A}')$；

(ii) 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是A的特征值，则

$$\sum_{i=1}^{n}|\lambda_i|^2\leqslant l$$

(上式是 Schur 不等式的改进式子)。

(提示：对那些 $B_k\neq 0$，$C_k\neq 0$ 的 B_k 与 C_k，令

$$\xi_{1k}^2=\mathrm{tr}(B_k\overline{B}_k')>0,\ \xi_{2k}^2=\mathrm{tr}(C_k\overline{C}_k')>0,$$

作方阵

$$K=\begin{pmatrix} A_k & \left(\dfrac{\xi_{2k}}{\xi_{1k}}\right)^{1/2}B_k \\ \left(\dfrac{\xi_{1k}}{\xi_{2k}}\right)^{1/2}C_k & D_k \end{pmatrix},$$

先证：$l_k=\mathrm{tr}(K\overline{K}')$，即得 (i)。再证$K$与$A$相似，说明$K$的 n 个特征值仍是 $\lambda_1, \lambda_2, \cdots, \lambda_n$，然后对$K$用 Schur 不等式，就可证得(ii)式。)

10. A, l_k, l 的假设同第 9 题，求证：

(i) $r(A)\geqslant\dfrac{1}{l}|\mathrm{tr}(A)|^2$，

(ii) 当 $|\mathrm{tr}(A)|^2>(n-1)l$，则 A 是非异阵。

(提示：完全仿照定理3.4的证法，并应用第9题的(ii)。)

11. 设 λ_1、$\lambda_2\cdots$，λ_n 是方阵 $A=(a_{ij})_{n\times n}$ 的特征值，记

$$d(A)=\max_{1\leqslant i,j\leqslant n}|\lambda_i-\lambda_j|,$$

称 $d(A)$ 为 A 的**展形**。又设 A 的分块与 l_k、l 的假设同第9题，求证：

(i) $d(A)\leqslant\sqrt{2}\left(l-\frac{1}{n}\left|\sum_{i=1}^{n}a_{ii}\right|^2\right)^{1/2}$；

(ii) 当 A 是奇异阵时，则

$$\max_{1\leqslant i\leqslant n}|\lambda_i|\leqslant\sqrt{2}\left(l-\frac{1}{n}\left|\sum_{i=1}^{n}a_{ii}\right|^2\right)^{1/2}.$$

并举例说明上面的等号可以达到。

(提示：运用第二章 §7 的 Lagrange 恒等式：

$$n\sum_{i=1}^{n}|\lambda_i|^2-\left|\sum_{i=1}^{n}\lambda_i\right|^2=\sum_{1\leqslant i<j\leqslant n}|\lambda_i-\lambda_j|^2.$$

先证出：

$$n\sum_{i=1}^{n}|\lambda_i|^2-\left|\sum_{i=1}^{n}\lambda_i\right|^2\geqslant\frac{n}{2}(d(A))^2$$

然后应用第9题的(ii)即得本题之(i)。)

12. A、l_k、l 的假设同第9题，

(i) 求证：

$$|\det A|\leqslant\frac{l^{n/2}}{n^{n/2}};$$

(ii) 设

$$D=\begin{vmatrix}3&3&2&2\\1&4&3&2\\1&-2&4&3\\-1&-3&-2&4\end{vmatrix},$$

求证：$|D|\leqslant 762$。

(提示：因为

$$|\det A|^2=|\lambda_1|^2|\lambda_2|^2\cdots|\lambda_n|^2,$$

此处 λ_1，λ_2，$\cdots\lambda_n$，是 A 的特征值，故对上式应用第6题提示中的 Cauchy 不等式，再应用第9题的(ii)即得(i)。由(i)即可估计(ii)的 D。)

13. 如果 n 阶复方阵 A 满足：$A\overline{A}'=\overline{A}'A$，则称 A 为**复正规阵**(当 A 是实方阵时，A 就是实正规阵)。证明：

(i) 复方阵 A 酉相似于对角阵的充要条件是，A 为复正规阵(这说明，除

了复正规阵外，再也没有其他复方阵可酉相似于复对角阵了）。

(ii) Schur 不等式的等号成立的充要条件是，A 为复正规阵。

(iii) 复正规阵必有完全特征向量系：$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，且

$$\bar{\alpha}_i'\alpha_j = \delta_{ij},$$

或者，

$$\alpha_i'\overline{\alpha_j} = \delta_{ij},\ i, j = 1, 2, \cdots, n.$$

14. 设 $A = (a_{ij})_{n\times n}$ 是复正规阵，记 $R(x) = |\bar{x}'Ax|$，其中 x 是满足 $\bar{x}'x = 1$（即长度为 1 的）n 维复的列向量，

(i) 求证：$\max\limits_{1\leqslant i\leqslant n}|\lambda_i| = \max\limits_{x} R(x)$，

此处 $\lambda_1, \lambda_2, \cdots\cdots, \lambda_n$ 是 A 的特征值。

(ii) 设 $|\lambda_1| = \max\limits_{1\leqslant i\leqslant n}|\lambda_i|$，求证

$$|\lambda_1| \geqslant \max_{1\leqslant i\leqslant n}|a_{ii}|$$

$$|\lambda_1| \geqslant \frac{1}{2}\max_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}}|a_{ii} + a_{jj} + a_{ij} + a_{ji}|.$$

（提示：应用复正规阵可酉相似于对角阵的结论先证出，

$$\max_{x} R(x) \leqslant \max_{1\leqslant i\leqslant n}|x_i|,$$

再应用 A 的谱分解式（§3 习题第 8 题）证出：

$$\max_{1\leqslant i\leqslant n}|\lambda_i| \leqslant \max_{x} R(x)).$$

15. 设 A 是复正规阵，记 $R(x, y) = |\overline{x'}Ax + \overline{y'}Ay|$，其中 x 与 y 是满足：$\bar{x}'x = 1, \bar{y}'y = 1, \bar{x}'y = 0$ 的 n 维复的列向量。又设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的特征值，且 $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$，求证：

$$|\lambda_1| + |\lambda_2| \geqslant \max_{x, y} R(x, y).$$

（提示：仿上题的证法；应用定理 3.2，并应用第二章 §7 中的 Cauchy-Schwarz 不等式。）

16. 设 A 是复正规阵，求证：必存在复系数多项式 $f(\lambda)$ 与 $g(\lambda)$，使

$$\overline{A'} = f(A),\ A = g(\overline{A'}).$$

（提示：设 A 的 n 个特征值是 $\lambda_1, \lambda_2, \cdots, \lambda_n$，不妨设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是其中所有的不相同的特征值。由定理 3.2，A 可写成

$$A = Q[\lambda_1, \lambda_2, \cdots, \lambda_n]\overline{Q'},$$

其中 Q 是酉阵，于是 $\overline{A'} = \overline{Q'}[\overline{\lambda_1}, \overline{\lambda_2}, \cdots, \overline{\lambda_n}]Q$，显然 $\overline{\lambda_1}, \overline{\lambda_2}, \cdots, \overline{\lambda_s}$ 也全不相同。于是作 Lagrange 内插多项式

$$g(\lambda)=\sum_{i=1}^{s}\frac{(\lambda-\bar{\lambda}_1)\cdots(\lambda-\bar{\lambda}_{i-1})(\lambda-\bar{\lambda}_{i+1})\cdots(\lambda-\bar{\lambda}_s)}{(\bar{\lambda}_i-\bar{\lambda}_1)\cdots(\bar{\lambda}_i-\bar{\lambda}_{i-1})(\bar{\lambda}_i-\bar{\lambda}_{i+1})\cdots(\bar{\lambda}_i-\bar{\lambda}_s)},$$

这个 $g(\lambda)$ 即可使 $A=g(\overline{A'})$。同法可找到 $f(\lambda)$，使 $\overline{A'}=f(A)$。）

17．(Fuglede-Putnam) 设 A 与 B 分别是 m 阶与 n 阶复正规阵，如果存在 $m\times n$ 阵 C，使 $AC=CB$，则 $\overline{A'}C=C\overline{B'}$。

（提示：由假设，$[A, B]$ 是 $m+n$ 阶复正规阵，再应用第 15 题的结论于 $[A, B]$ 即可得证。）

18．(Wiegmann, 1970) 设 A 与 B 分别是 $m\times n$ 复方阵与 $n\times m$ 复方阵，则 AB 与 BA 都是复正规阵的充要条件是，

$$\overline{A'}AB=BA\overline{A'},\qquad \overline{B'}BA=AB\overline{B'}.$$

（提示：应用第 16 题的结论。）

19．设 $c_1, c_2, \cdots, c_{n-1}$ 不全为零，称方阵

$$\begin{pmatrix} h_{11} & h_{12} & h_{13}\cdots h_{1,n-1} & h_{1n} \\ c_1 & h_{22} & h_{23}\cdots h_{2,n-1} & h_{2n} \\ & c_2 & \ddots \quad \vdots & \vdots \\ & & \ddots \; h_{n-1,n-1} & h_{n-1,n} \\ & 0 & c_{n-1} & h_{nn} \end{pmatrix}=\begin{pmatrix} \diagdown\!\diagdown & * \\ 0 & \diagdown\!\diagdown \end{pmatrix}$$

为上 Hessenberg 阵，而称

$$\begin{pmatrix} h_{11} & c_1 & & & \\ h_{21} & h_{22} & c_2 & & \\ h_{31} & h_{32} & h_{33} & \ddots & \\ \vdots & \vdots & \vdots & \ddots & c_{n-1} \\ h_{n1} & h_{n2} & h_{n3} & \cdots & h_{nn} \end{pmatrix}=\begin{pmatrix} \diagdown\!\diagdown & 0 \\ * & \diagdown\!\diagdown \end{pmatrix}$$

为下Hessenberg 阵（例如 Frobenius 块是上 Hessenberg 阵，Jacobson 块是下 Hessenberg 阵）。求证：任何 n 阶复方阵 $a=(a_{ij})_{n\times n}$ 必酉相似于上 Hessenberg 阵，即存在酉阵 Q，使

$$\overline{Q'}AQ=\begin{pmatrix} \diagdown\!\diagdown & * \\ 0 & \diagdown\!\diagdown \end{pmatrix}.$$

（提示：仿定理 3.1 的证法，把 A 分块：

$$A=\begin{pmatrix} a_{11} & \beta_1 \\ \alpha_1 & A_1 \end{pmatrix}$$

设 $\|\alpha_1\|=C_1$，应用镜象阵的基本定理，可找到 $n-1$ 阶复镜象阵 H_1，使

$$H_1\alpha_1=(C_1, 0, \cdots\cdots, 0)'.$$

令 $Q_1=[1, H_1]$，则 Q_1 是 n 阶复镜象阵，且

$$Q_1A\overline{Q_1'}=\begin{pmatrix} a_{11} & \beta_1\overline{H_1'} \\ H_1\alpha_1 & H_1A\overline{H_1'} \end{pmatrix}=\left(\begin{array}{c|c} a_{11} & \beta_1\overline{H_1'} \\ \hline c_1 & \\ 0 & H_1A\overline{H_1'} \\ \vdots & \\ 0 & \end{array}\right)$$

$$=\left(\begin{array}{cc|c} a_{11} & b_{12} & B_2 \\ c_1 & b_{22} & \\ \hline 0 & \alpha_2 & A_2 \end{array}\right),$$

然后对 α_2 与 $n-2$ 阶阵 A_2 用归纳法。)

20．求证：对任何实方阵 A，必可找到正交阵 Q，使 $Q'AQ$ 是实的上 Hessenberg 阵。

21．对任一 n 阶 Hermite 阵，求证：必存在酉阵 Q，使 $\overline{Q'}AQ$ 是三对角 Hermite 阵，即，

$$\overline{Q'}AQ=\begin{pmatrix} a_1 & \bar{b}_1 & & & 0 \\ b_1 & a_2 & \bar{b}_2 & & \\ & b_2 & a_3 & \ddots & \\ & 0 & \ddots & \ddots & \bar{b}_{n-1} \\ & & & b_{n-1} & a_n \end{pmatrix},$$

其中 a_i, b_i 全是实数，$\bar{b}_j$ 是 b_j 的共轭复数，

$$i=1, 2, \cdots, n; j=1, 2, \cdots, n-1.$$

22．求证：任何 n 阶实对称阵必正交相似于三对角实对称阵，

$$\begin{pmatrix} a_1 & b_1 & & & 0 \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & \ddots & \\ 0 & & \ddots & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_n \end{pmatrix}$$

此处 a_i 全是实数，$i=1, 2, \cdots, n$。

23．设 A、B、C 都是 n 阶实方阵，且存在正交阵 Q，使

$$Q'BQ=(h_{ij})_{n\times n}=H,$$

H 是上 Hessenberg 阵(即 $h_{ij}=0$, 当 $i>j+2$ 时), 又记

$$f(\lambda)=|\lambda I_n+H|=|\lambda I_n+B|.$$

令 $\lambda\delta_{ij}+h_{ij}$ (在 $f(\lambda)$ 中)的代数余子式是 $f_{ij}(\lambda)$, $i, j=1, 2, \cdots, n$。求证:

(i) 矩阵方程 $AX+XB'=C$ 有唯一解的充要条件是 $f(A)$ 为非异阵。

(ii) 当 $f(A)$ 是非异阵, 且 $\prod\limits_{j=2}^{n} h_{j, j-1}\neq 0$ 时, 记

$$A_1=Q'AQ,\ Q'CQ=(\delta_1, \delta_2, \cdots, \delta_n),$$

则 $AX+XB'=C$ 的唯一解可以由解:

$$X=Q(\alpha_1, \alpha_2, \cdots, \alpha_n)Q',$$

$$f(A_1)\alpha_n=\sum_{i=1}^{n} f_{in}(A_1)\delta_i,$$

以及下面的以列向量 $\alpha_1, \alpha_2, \cdots, \alpha_{n-1}$ 为未知向量的上三角形向量方程组:

$$\sum_{j=i-1}^{n-1}(A_1\delta_{ij}+h_{ij}I_n)\alpha_j=\beta_i,\ i=2, 3, \cdots, n$$

而得出, 其中 $\beta_i=\delta_i-h_{in}\alpha_n$, $i=2, \cdots, n-1$; $\beta_n=\delta_n-(A_1+h_{nn}I_n)\alpha_n$。

(注: 由(ii)可知, 只要解出 X 的 n^2 个未知量中的 n 个, 即解出 α_n, 则 X 的其他 n^2-n 个未知量几乎立刻可以解出, 此法对较大的 n 更有用, 这正好说明方阵正交相似的作用。)

(提示: 首先把矩阵方程 $AX+XB'=C$ 化成同解的矩阵方程:

$$A_1(Q'XQ)+(Q'XQ)H'=Q'CQ,$$

由假设条件, 上方程又可化为同解的向量方程组:

$$\sum_{j=1}^{n}(\delta_{ij}A_1+h_{ij}I_n)\alpha_j=\delta_i,\ i=1, 2, \cdots, n, \tag{$\triangle$}$$

再由行列式的展开式以及 Cramer 法则及其逆定理, 即可证得 (i)。又以 $f_{in}(A_1)$ 依次左乘($\triangle$)的第 i 个方程, $i=1, 2, \cdots, n$, 并经过简单的计算, 即可证得(ii)。)

24. 设 A、C 都是 n 阶实方阵,

$$g(\lambda)=\frac{1}{2}\{|\lambda I_n+A|+|\lambda I_n-A|\},$$

$$l(\lambda)=\frac{1}{2}\{|\lambda I_n+A|-|\lambda I_n-A|\},$$

求证: 矩阵方程 $AX+XA'=C$ 有唯一解的充要条件是, $g(A)$ (或 $l(A)$) 为非异阵。

(提示: 应用第 19 题 (i) 以及 Hamilton-Cayley 定理。)

第八章　方阵的合同与二次型

本章详细介绍实二次型化为标准形的问题，重点讨论正定二次型、半正定二次型以及与之相应的正定阵与半正定阵的各种基本性质及其应用，特别对正定阵作了较为细致的讨论.

§1　二次型的简化问题、方阵的合同

一、二次型的简化问题

二次型的理论起源于二次曲线(或曲面)的简化问题. 我们知道，平面上的二次曲线方程:

$$ax^2+bxy+cy^2+dx+ey=f \tag{1}$$

可用坐标变换

$$\begin{cases} x=x'\cos\theta-y'\sin\theta \\ y=x'\sin\theta+y'\cos\theta \end{cases} \tag{2}$$

把(1)化成

$$a'x'^2+c'y'^2+d'x'+e'y'=f. \tag{3}$$

由(3)就可决定二次曲线的形状，所以对(1)来说，关键在于用坐标变换(2)把(1)中二次齐次项:

$$f(x,y)=ax^2+bxy+cy^2 \tag{4}$$

的非平方项消掉. 对二次曲面也可类似地这样做.

在数学分析中. 判断 $f(x,y)$ 的某一极值可疑点 (a,b) 是否为 $f(x,y)$ 的极大点或极小点，是用下法解决的：如果函数 $f(x,y)$ 具有各阶连续偏导数，则先把 $f(x,y)$ 写成 Taylor 展开式:

$$f(x,y)=f(a,b)+\left.\frac{\partial f}{\partial x}\right|_{x=a,y=b}(x-a)+\left.\frac{\partial f}{\partial y}\right|_{x=a,y=b}(y-b)$$

$$+\frac{\partial^2 f}{\partial x^2}\bigg|_{x=a,y=b}\frac{(x-a)^2}{2!}+\frac{\partial^2 f}{\partial x\partial y}\bigg|_{x=a,y=b}$$

$$\cdot(x-a)(y-b)+\frac{\partial^2 f}{\partial y^2}\bigg|_{x=a,y=b}\frac{(y-b)^2}{2!}+R,$$

其中R是余项。由于(a, b)是$f(x, y)$的极值可疑点，故

$$\frac{\partial f}{\partial x}\bigg|_{x=a,y=b}=0,\ \frac{\partial f}{\partial y}\bigg|_{x=a,y=b}=0,$$

于是

$$f(x, y)=f(a, b)+\frac{1}{2!}\cdot\frac{\partial^2 f}{\partial x^2}\bigg|_{x=a,y=b}(x-a)^2+\frac{\partial^2 f}{\partial x\partial y}\bigg|_{x=a,y=b}$$

$$\cdot(x-a)(y-b)+\frac{1}{2!}\frac{\partial^2 f}{\partial y^2}\bigg|_{x=a,y=b}(y-b)^2+R.$$

记

$$x-a=z_1,\ y-b=z_2;$$

$$\frac{1}{2!}\cdot\frac{\partial^2 f}{\partial x^2}\bigg|_{x=a,y=b}=a_{11},\ \frac{1}{2!}\frac{\partial^2 f}{\partial x\partial y}\bigg|_{x=a,y=b}=a_{12},$$

$$\frac{1}{2!}\cdot\frac{\partial^2 f}{\partial y^2}\bigg|_{x=a,y=b}=a_{22},$$

则判断(a, b)是极大点，还是极小点，就是要看$f(x, y)$的二次项：

$$g(z_1, z_2)=a_{11}z_1^2+2a_{12}z_1z_2+a_{22}z_2^2 \tag{5}$$

在区域：$|x-a|<\epsilon$，$|y-b|<\epsilon$ 上小于零，还是大于零（ϵ 是充分小的正数）。

容易看出，如果能用类似于(2)式的坐标变换把(5)式右边的非平方项消掉，则即可判断(a, b)是极大点，还是极小点。

上面两例表明，用一个坐标变换把二次型（二次齐次多项式）$ax^2+bxy+cy^2$的非平方项消掉乃是必要的。在力学、物理学以及一些数学分支中也需要用坐标变换（称为非异坐标变换）：

$$x_i=\sum^n q_{ij}y_j,\ i=1, 2, \cdots, n,\ |q_{ij}|_{n\times n}\neq 0, \tag{6}$$

把二次型 $\sum\limits_{i,j=1}^n a_{ij}x_ix_j$ 的非平方项消掉，即把(6)式代入

$$\sum_{i,j=1}^n a_{ij}x_ix_j$$

使之化成

$$\sum_{i,j=1}^{n} a_{ij}x_ix_j = \sum_{i=1}^{n} d_iy_i^2 \tag{7}$$

的形状．这里 x_i，y_i 都是实变数；d_i，q_{ij}，a_{ij} 都是实数，i、$j=1$、2，…，n。这个问题就称作(实)二次型的简化问题，或称为把二次型化为平方和。

若记

$$x=(x_1, x_2, \cdots, x_n)', \; y=(y_1, y_2, \cdots, y_n)',$$

$$A=(a_{ij})_{n\times n}, \; Q=(q_{ij})_{n\times n}, \; D=[d_1, d_2, \cdots, d_n],$$

则二次型的简化问题就归结为找非异坐标变换：$x=Qy$、$|Q|\neq 0$，使

$$\sum_{i,j=1}^{n} a_{ij}x_ix_j = x'Ax = y'Dy = \sum_{i=1}^{n} d_iy_i^2. \tag{8}$$

因为任意 n 阶实方阵 A 必有唯一分解式：

$$A=B+C,$$

其中 $B=\dfrac{A+A'}{2}$ 是实对称阵，$C=\dfrac{A-A'}{2}$ 是反对称阵，所以

$$x'Ax=x'(B+C)x=x'Bx$$

(因为 $x'Cx=0$)。这就是说，任何实二次型 $x'Ax$ 可化成以 B 为实对称阵的二次型 $x'Bx$。所以今后就假设 A 是实对称阵，称 A 为二次型 $x'Ax$ 的**系数矩阵**。以下讨论的(实)二次型都是以实对称阵为系数矩阵的二次型，不再另作说明。

若将 $x=Qy$ 代入 $x'Ax$，则得到

$$x'Ax=y'(Q'AQ)y, \tag{9}$$

所以只要 $Q'AQ$ 是对角阵 D：$Q'AQ=D$，则即可得到(8)式。于是化二次型为平方和的问题就归结为对于实对称阵 A 怎样找非异阵 Q，使

$$Q'AQ=[d_1, d_2, \cdots, d_n]=D, \tag{10}$$

这就归结为方阵的简化问题。于是引进

定义　设 A 与 B 都是同阶实方阵，称 A 与 B 是**合同**的，或称 A **合同于** B，如果存在非异实方阵 Q，使 $Q'AQ=B$。

由方阵间合同的定义可知，化二次型为平方和的问题就归结为把实对称阵合同于对角阵的问题。

如果Q是正交阵，则$Q'AQ=Q^{-1}AQ$，此时合同与相似是一致的。

二、用初等变换化二次型为平方和

第七章定理 2.2 指出，对任何n阶实对称阵A，必可找到正交阵Q，使$Q'AQ=[\lambda_1, \lambda_2, \cdots, \lambda_n]$，而$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的特征值。用化二次型为平方和的术语来描述，就是下面的

定理 1.1 必可找到“正交坐标变换”：$x=Qy$，Q是正交阵，把二次型$x'Ax$化为平方和$y'[\lambda_1, \lambda_2, \cdots, \lambda_n]y$，即

$$x'Ax=\sum_{i,j=1}^{n}a_{ij}x_ix_j=\sum_{i=1}^{n}\lambda_iy_i^2,$$

而$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的特征值。

用正交坐标变换把二次型化为平方和，常称为化二次型到主轴上去。这就是第七章定理 2.2 的“几何”背景。

今问，对任意数域K，是否存在非异阵Q，使$Q'AQ$为K上对角阵呢?回答是肯定的，为此先证：

引理 1 设A是n阶实对称阵，A_r是A的r阶非异顺序主子阵，A的分块是

$$A=\begin{pmatrix}A_r & B\\ C & D\end{pmatrix},$$

则必：

(i) $A_r=A_r'$, $D=D'$, $C=B'$;

(ii) 必存在非异阵Q，使

$$Q'AQ=[A_r, D-B'A_r^{-1}B]. \tag{11}$$

【证明】 (i)是明显的。今证(ii)。对分块阵A作第三种分块初等阵的列变换及其相应的转置行变换，即只要取非异阵Q'为：

$$Q'=T_{21}(-B'A_r^{-1})=\begin{pmatrix}I_r & 0\\ -B'A_r^{-1} & I_{n-r}\end{pmatrix},$$

就使(11)式成立。

定理 1.2 设A是秩为r的n阶实对称阵，$r\geq 1$，则必存在n阶非异实方阵Q（Q不一定是正交阵），使

$$Q'AQ=[d_1,\ d_2,\ \cdots,\ d_r,\ 0,\ \cdots,\ 0], \tag{12}$$

其中$d_1,\ d_2,\ \cdots,\ d_r$都是非零实数。

【证明】 下面是个构造性的证明。对n用归纳法。$n=1$时定理显然正确，设定理对$n-1$成立，

(1) 若A的$(1,\ 1)$元$a_{11}\neq 0$，把A如下分块：

$$A=\begin{pmatrix}a_{11} & \alpha \\ \alpha' & A_1\end{pmatrix}.$$

由引理1，必存在非异阵Q_1，使

$$Q_1'AQ_1=[a_{11},\ A_1-\alpha' a_{11}^{-1}\alpha]. \tag{13}$$

由于A_1也是对称阵，所以$A_1-\alpha' a_{11}^{-1}\alpha$仍是$n-1$阶对称阵，且由(13)式可知，$A_1-\alpha' a_{11}^{-1}\alpha$的秩为$r-1$，于是由归纳法假设，存在$n-1$阶非异阵$Q_2$，使

$$Q_2'(A_1-\alpha' a_{11}^{-1}\alpha)Q_2=[d_2,\ \cdots,\ d_r,\ 0,\ \cdots,\ 0], \tag{14}$$

而$d_2,\ \cdots,\ d_r$都是非零实数。记$Q=Q_1[1,\ Q_2]$，则Q是n阶非异阵。故由(13)式与(14)式並经过简单的计算，即可证得(12)式（其中$a_{11}=d_1$）。

(2) 当$a_{11}=0$，而某一$a_{ii}\neq 0(i>1)$，则

$$E_{1i}'AE_{1i}=\begin{pmatrix}a_{ii} & * \\ * & B\end{pmatrix},$$

其中E_{1i}是第一种初等阵。于是又化为第(1)种情形。

(3) 如果$a_{ii}=0,\ i=1,2,\cdots,n$，则至少有一个$a_{ij}\neq 0(i\neq j)$否则$A=O$，于是作第三种初等变换$T_{ij}(a_{ij})$及其转置变换$T_{ji}(a_{ij})$，即得

$$T_{ji}'(a_{ij})AT_{ji}(a_{ij})=\begin{pmatrix}0 & \cdots & \cdots & \cdots & \cdots & * \\ & \ddots & & & & \\ \cdots & 2a_{ij} & \cdots & a_{ij} & \cdots & \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \cdots & a_{ij} & \cdots & 0 & \cdots & \\ \vdots & \vdots & & \cdots & \ddots & \cdots \\ * & \cdots & \cdots & \cdots & \cdots & 0\end{pmatrix},$$

于是位于上式右边的 (i, i) 元素 $2a_{ij} \neq 0$，这样又化成了(1)的情形。 证毕。

如果称二次型 $x'Ax$ 的系数矩阵的秩为二次型的秩，则可把定理1.2改述为

定理 1.2' 必可找到非异坐标变换 $x = Qy$，把秩为 $r > 0$ 的二次型 $x'Ax$ 化成平方和：

$$x'Ax = d_1y_1^2 + d_2y_2^2 + \cdots + d_ry_r^2 (= y'[d_1, d_2, \cdots, d_r, 0, \cdots, 0]y), \quad (15)$$

而 $d_1, d_2, \cdots, d_r$ 全不为零。

称(15)式右端的平方和为二次型 $x'Ax$ 的标准形。

在第一章§5中已找到了求 Q 的方法，即作下列有限次初等变换：

$$\begin{pmatrix} A & I \\ I & 0 \end{pmatrix} \to \cdots \to \begin{pmatrix} Q'AQ & Q' \\ Q & 0 \end{pmatrix},$$

在找出 $Q'AQ$ 为对角阵的同时，也求得了 Q。对于实对称阵来说，用下面两个引理还可使 Q 的求法进一步简化：

引理 2 对任何秩为 r 的 n 阶实对称阵 A，必可找到"正么模阵"B，即 B 是有限个第三种初等阵 $T_{ij}(k)$ 的乘积(显然 $|B| = 1$)，使

$$B'AB = [d_1, d_2, \cdots, d_r, 0, \cdots, 0], \quad (16)$$

而 $d_1, d_2, \cdots, d_r$ 全不为零。

【证明】 由定理1.2，存在 n 阶非异阵 Q，使

$$Q'AQ = [c_1, c_2, \cdots, c_r, 0, \cdots, 0], \quad (17)$$

而 $c_i \neq 0$，$i = 1, 2, \cdots, r$。由第二章定理，

$$Q = BD_n(\mu), \ \mu = |Q| \neq 0, \quad (18)$$

其中 B 是有限个第三种初等阵的乘积，而 $D_n(\mu) = [1, \cdots, 1, \mu]$。把(18)式代入(17)式，得到：

$$\begin{aligned} B'AB &= D_n(\mu^{-1})[c_1, c_2, \cdots, c_r, 0, \cdots, 0]D_n(\mu^{-1}) \\ &= [d_1, d_2, \cdots, d_r, 0, \cdots, 0], \end{aligned}$$

其中 $d_i = c_i$，$i = 1, 2, \cdots, r-1$；而 $d_r = c_r$，当 $r < n$；$d_r = \mu^{-2}c_n$，当 $r = n$。故 $d_1, d_2, \cdots, d_r$ 全不为零。 证毕。

引理2说明，必可用有限个第三种初等阵把实对称阵合同于一个对角阵。

引理3 设A是n阶实对称阵，如果Q'是有限个第三种初等阵$T_{i1}(k)(i>1)$的乘积，且使

$$Q'A=\begin{pmatrix} d_1 & \delta \\ 0 & A_1 \end{pmatrix},$$

其中δ是$n-1$维行向量，A_1是$n-1$阶阵，则必

$$Q'AQ=\begin{pmatrix} d_1 & 0 \\ 0 & A_1 \end{pmatrix}.$$

【证明】 由假设，Q'是$T_{i1}(k)$的乘积，$i>1$，故Q是$T_{1i}(k)$的乘积，$i>1$。由于$i>1$，故$Q'A$左乘Q时，对$Q'A$的第1列的元素没有影响（因$BT_{1i}(k)$是把B的第1列的k倍加到B的其他列上去，而B的第1列不动），而变动的只是δ的元素，故得

$$Q'AQ=\begin{pmatrix} d_1 & \beta \\ 0 & A_1 \end{pmatrix},$$

但A是对称阵，故$Q'AQ$也是对称阵，所以$\beta=0$。 证毕。

由引理3，只要找Q'使

$$Q'(A,I)=(Q'A,Q')=\left(\begin{pmatrix} d_1 & & & & & \\ & d_2 & & & * & \\ & & \ddots & & & \\ & & & d_r & & \\ & & & & 0 & \\ & 0 & & & & \ddots \\ & & & & & & 0 \end{pmatrix}, Q'\right), \tag{19}$$

则这个Q'的转置阵就是要找的非异阵Q，它使$Q'AQ$为对角阵。换言之，只要对(A,I)作有限次第三种初等变换$T_{ij}(k)$，$i>j$，那么，当把A变成上三角阵时，(A,I)的I就同时化为Q'，它的转置阵就是要找的非异阵。

例1 求非异阵Q，使$Q'AQ$为对角阵，其中

$$A=\begin{pmatrix} 1 & -1 & 2 \\ -1 & -1 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

【解】 逐次用第三种初等行变换．$T_{ij}(k)$，$i>j$，把 (A, I_3) 化成(19)式的形状：

$$(A,I_3)=\left(\begin{array}{ccc|ccc}1&-1&2&1&0&0\\-1&-1&0&0&1&0\\2&0&2&0&0&1\end{array}\right)\to\left(\begin{array}{ccc|ccc}1&-1&2&1&0&0\\0&-2&2&1&1&0\\2&0&2&0&0&1\end{array}\right)\to$$

$$\to\left(\begin{array}{ccc|ccc}1&-1&2&1&0&0\\0&-2&2&1&1&0\\0&2&-2&-2&0&1\end{array}\right)\to\left(\begin{array}{ccc|ccc}1&-1&2&1&0&0\\0&-2&2&1&1&0\\0&0&0&-1&1&1\end{array}\right),$$

故由引理 3 可知

$$Q'AQ=[1,\ -2,\ 0],$$

而

$$Q=(Q')'=\begin{pmatrix}1&0&0\\1&1&0\\-1&1&1\end{pmatrix}'=\begin{pmatrix}1&1&-1\\0&1&1\\0&0&1\end{pmatrix}.$$

例 2 把二次型 $2x_1x_2-6x_2x_3+2x_1x_3$ 化为平方和。

【解】 由于二次型的系数矩阵是

$$A=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix},$$

而 A 的主对角元全是零，所以不能立即用引理 3，先由引理 2 以及定理 1.2 证明过程中的 (3)，对 A 作第三种初等行变换及其转置列变换，使得经过如此变换后得到的新的合同阵的主对元有非零数，然后对这个矩阵用引理 3，整个变换过程如下：

$$(A,I_3)=\left(\begin{array}{ccc|ccc}0&1&1&1&0&0\\1&0&-3&0&1&0\\1&-3&0&0&0&1\end{array}\right)\to\left(\begin{array}{ccc|ccc}1&1&-2&1&1&0\\1&0&-3&0&1&0\\1&-3&0&0&0&1\end{array}\right)\to$$

(此步不影响竖线右方的矩阵)

$$\xrightarrow{} \overset{\left(-\frac{1}{2}\right)}{\longrightarrow}\begin{pmatrix} 2 & 1 & -2 & 1 & 1 & 0 \\ 1 & 0 & -3 & 0 & 1 & 0 \\ -2 & -3 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{(1)} \begin{pmatrix} 2 & 1 & -2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -2 & -3 & 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow$$

$$\xrightarrow{(-4)} \begin{pmatrix} 2 & 1 & -2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -2 & -2 & 1 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & -2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 6 & 3 & -1 & 1 \end{pmatrix},$$

所以

$$Q = \begin{pmatrix} 1 & -\frac{1}{2} & 3 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

而

$$Q'AQ = [2,\ -\frac{1}{2}, 6], \tag{20}$$

故二次型 $2x_1x_2-6x_2x_3+2x_1x_3$ 的标准形是

$$2y_1^2-\frac{1}{2}y_2^2+6y_3^2. \tag{21}$$

而非异坐标变换是

$$\begin{cases} x_1 = y_1-\frac{1}{2}y_2+3y_3 \\ x_2 = y_1+\frac{1}{2}y_2-y_3 \\ x_3 = y_3 \end{cases}$$

它把原二次型化为形如(21)式的平方和。

必须注意：

(i) 上述 Q' 只是第三种初等阵 $T_{ij}(k)$ 的乘积，但未必有 $i>j$，因为一开始已作了 $T_{12}(1)$ 的行变换；

(ii) 非异阵 Q 不是唯一的。例如由(20)式还可得：

$$(QD_2(-2))'A(QD_2(-2)) = [1,\ -2,\ 1]Q'AQ[1,\ -2,\ 1] = [2,\ -2,\ 6],$$

故得

$$Q_1'AQ_1=[2,\ -2,\ 6].$$

而

$$Q_1=QD_2(-2)=\begin{pmatrix}1&1&3\\1&-1&-1\\0&0&1\end{pmatrix},\qquad(22)$$

故原二次型的另一个标准形是

$$2z_1^2-2z_2^2+6z_3^2.\qquad(23)$$

用初等变换把二次型 $x'Ax$ 化为平方和的最大优点在于，化 A 为上三角阵的同时就求出了 Q。

三、配方法(简述)

把二次型化为平方和的另一方法就是在初等代数里已熟悉的配(平)方法，此法称为 Lagrange 方法。此法的缺点是，不能在化 $Q'AQ$ 为对角阵的同时就找出 Q，而要经过多次的方阵的乘法运算后才能得出 Q，因而计算复杂而费时，因而是不实用的。所以我们不准备对一般的配方法进行讨论，然而在有些理论推导中以及对某些形状特殊的二次型，用配方法却容易奏效。故本段仍举例以说明此法的运用过程。

仍以例 2 为例。因为二次型

$$f(x_1,\ x_2,\ x_3)=2x_1x_2-6x_2x_3+2x_1x_3$$

没有平方项，故无从进行配方，所以先作非异坐标变换：

$$x_1=z_1+z_2$$
$$x_2=z_1-z_2$$
$$x_3=z_3$$

即作

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix},\qquad(24)$$

于是

$$f(x_1, x_2, x_3) = 2z_1^2 - 2z_2^2 - 4z_1z_3 + 8z_2z_3,$$

上式已出现平方项，即可进行配方：

$$\begin{aligned} f(x_1, x_2, x_3) &= 2(z_1 - z_3)^2 - 2z_2^2 - 2z_3^2 + 8z_2z_3 \\ &= 2y_1^2 - 2y_2^2 - 2y_3^2 + 8y_2y_3, \end{aligned} \tag{25}$$

此处

$$\begin{aligned} y_1 &= z_1 - z_3 \\ y_2 &= z_2 \\ y_3 &= z_3 \end{aligned}$$

或者

$$\begin{aligned} z_1 &= y_1 + y_3 \\ z_2 &= y_2 \\ z_3 &= y_3 \end{aligned}$$

也即

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \tag{26}$$

再对(25)式进行配方：

$$\begin{aligned} f(x_1, x_2, x_3) &= 2y_1^2 - 2(y_2 - 2y_3)^2 + 6y_3^2 \\ &= 2u_1^2 - 2u_2^2 + 6u_3^2, \end{aligned} \tag{27}$$

此处

$$\begin{aligned} y_1 &= u_1 \\ y_2 &= u_2 + 2u_3 \\ y_3 &= u_3 \end{aligned}$$

也即

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \tag{28}$$

由(24)、(26)、(28)诸式即得，

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}$$

$$=\begin{pmatrix}1&1&3\\1&-1&-1\\0&0&1\end{pmatrix}\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix},$$

故

$$Q=\begin{pmatrix}1&1&3\\1&-1&-1\\0&0&1\end{pmatrix},$$

使

$$Q'AQ=[2,\ -2,\ 6]$$

(与(22)、(23)式相比较)。

从上面的配方过程可以看出,Q 不仅需要另外求出,而且要在对逐次配方后所得的非异阵求出其逆后,才能逐个相乘得出 Q,其麻烦程度可想而知。然而,用配方法求某些特殊二次型的标准形却很有效,下例便是。

例 3 求 $x=Qy$, 化二次型

$$f(x_1,x_2,\cdots,x_{2n})=x_1x_{2n}+x_2x_{2n-1}+\cdots+x_nx_{n+1}$$

为标准形。其中 $x=(x_1,x_2,\cdots,x_{2n})'$, $y=(y_1,y_2,\cdots,y_{2n})'$。

【解】 用配方法。令

$$\begin{aligned}x_i&=y_i+y_s,\\x_s&=y_i-y_s,\end{aligned}\quad i=1,2,\cdots,n;\ s=2n,2n-1,\cdots,n+1,\tag{29}$$

即

$$x=\begin{pmatrix}x_1\\\vdots\\x_n\\\hline x_{n+1}\\\vdots\\x_{2n}\end{pmatrix}=\begin{pmatrix}I_n&J_n\\J_n&-I_n\end{pmatrix}\begin{pmatrix}y_1\\\vdots\\y_n\\\hline y_{n+1}\\\vdots\\y_{2n}\end{pmatrix}=Qy,\tag{30}$$

其中 $J_n=(e_n,\cdots,e_2,e_1)$, e_i 是标准单位向量,$1\leqslant i\leqslant n$, 而

$$Q=\begin{pmatrix} I_n & J_n \\ J_n & -I_n \end{pmatrix}$$

显然是非异阵(由行列式的第一降阶定理立即可以看出)。把(29)代入 $f(x_1, x_2, \cdots, x_{2n})$，就得到了它的标准形：

$$f(x_1, x_2, \cdots, x_{2n})=\sum_{i=1}^{n} x_i x_{2n-i+1}=\sum_{i=1}^{n}(y_i^2-y_{n-i+1}^2)$$

$$=\sum_{i=1}^{n} y_i^2-\sum_{i=n+1}^{2n} y_i^2.$$

习　题

1. 求非异坐标变换 $x=Qy$，把下列二次型化为标准形：

(i) $x_1^2+2x_1x_2+4x_1x_3-x_3^2$；

(ii) $2x_1x_2+6x_2x_3+2x_1x_3$；

(iii) $x_1^2-4x_1x_2+2x_1x_3-2x_1x_4+3x_2^2-6x_2x_3+8x_2x_4+2x_3x_4-3x_4^2$；

(iv) $\sum_{i=1}^{n-1} x_i x_{i+1}$；

(v) $\sum_{i=1}^{n} x_i^2+\sum_{1\leqslant i<j\leqslant n} x_i x_j$；

(vi) $\sum_{i=1}^{n}(x_i-x_M)^2$，此处 $x_M=\dfrac{1}{n}\sum_{i=1}^{n} x_i$。

2. (i) 求证：合同的方阵集合具有自反性(即 A 与 A 本身合同)、对称性(即若 A 与 B 合同，则 B 与 A 合同)、传递性(即若 A 与 B 合同，B 与 C 合同，则 A 与 C 合同)。

(ii) 把 n 阶实对称阵集合 S 中凡合同的所有阵归入同一类，称为"合同类"，问 S 可分成多少不同的合同类?

3. 设 A 是 n 阶实对称阵，求证：必存在正实数 k，使对任一 n 维实的列向量 α，恒有 $|\alpha' A\alpha|\leqslant k\alpha'\alpha$。

4. 求证：任何秩为 r 的 n 阶实对称阵 A 必有分解式：$A=T'T$，而 T 是秩为 r 的 n 阶复方阵。

5. 对(实)二次型 $x'Ax$，如存在 n 维实的列向量 α 与 β，使 $\alpha' A\alpha<0$，$\beta' A\beta>0$，求证：必存在 n 维实的列向量 δ，使 $\delta' A\delta=0$。

6. 设 A 是秩为 r 的 n 阶实对称阵，求证：必可找到与 A 合同的实对称 B，使 $\left|B\begin{pmatrix}1\\1\end{pmatrix}\right|$，$\left|B\begin{pmatrix}1&2\\1&2\end{pmatrix}\right|$，…，$\left|B\begin{pmatrix}12\cdots r\\12\cdots r\end{pmatrix}\right|$ 中任意两个相邻的顺序主子

式不全为零。并且，当 $\left|B\begin{pmatrix}12\cdots i\\12\cdots i\end{pmatrix}\right|=0$时，$i<r$ 时，则 $\left|B\begin{pmatrix}12\cdots i-1\\12\cdots i-1\end{pmatrix}\right|$ 与 $\left|B\begin{pmatrix}12\cdots i+1\\12\cdots i+1\end{pmatrix}\right|$ 的符号相反。

（提示：应用第三章选做题第 9 题及第二章 §5 例 7.）

7. 设(实)二次型 $f(x)=f(x_1, x_2, \cdots, x_n)=\sum\limits_{i,j=1}^{n} a_{ij}x_ix_j$，求证：

(i) 如果 $a_{ii}\neq 0$，则 $f(x)=\frac{1}{4a_{ii}}\left(\frac{\partial f}{\partial x_i}\right)^2+f_1(x)$，此处 $f_1(x)$是不含 x_i 的(实)二次型，$1\leqslant i\leqslant n$。

(ii) 如果 $a_{ii}=0$，$a_{jj}=0$，但 $a_{ij}\neq 0$，则

$$f(x)=\frac{1}{8a_{ij}}\left[\left(\frac{\partial f}{\partial x_i}+\frac{\partial f}{\partial x_j}\right)^2-\left(\frac{\partial f}{\partial x_i}-\frac{\partial f}{\partial x_j}\right)^2\right]+f_2(x),$$

此处 $f_2(x)$为不含 x_i 与 x_j 的实二次型，$1\leqslant i, j\leqslant n$。

8. 应用第 8 题把(实)二次型：

$$4x_1^2+x_2^2+x_3^2+x_4^2-4x_1x_2-4x_1x_3+4x_1x_4+4x_2x_3-4x_3x_4$$

化为平方和(不要求找出合同变换矩阵 Q)。

9. 设

$$M=\begin{pmatrix}A & B\\-B' & D\end{pmatrix}$$

是实方阵，且 A、D 都是方阵，

(i) 如果A是非异阵，求证：必存在非异阵P与 Q，使

$$PMQ=[A,\ D+B'A^{-1}B];$$

(ii) 如果M是反对称(实)方阵，A 是非异阵，求证：A 与 D、$D+B'A^{-1}B$ 都是反对称(实)方阵，且存在非异阵 Q，使

$$Q'AQ=[A,\ D+B'A^{-1}B];$$

(iii) 对任何反对称(实)方阵 A，求证：必存在非异阵 Q，使得

$$Q'AQ=\left[0,\ \begin{pmatrix}0&1\\-1&0\end{pmatrix},\ \begin{pmatrix}0&1\\-1&0\end{pmatrix},\ \cdots,\ \begin{pmatrix}0&1\\-1&0\end{pmatrix}\right].$$

§ 2　惯性定律、二次型的分类

一、惯性定律

上节例 2 的(21)式与(23)式表明，可以用不同的非异坐标变

换把二次型化成不同的标准形，然而标准形(21)与(23)也有共同之处，即它们的正项个数相同（都是 2 个）、负项个数相同（都是 1 个）。这是很必要的，因为二次曲线或二次曲面(的形状)不会因为坐标系的改变而改变，所以对含有 n 个实变量 $x_1, x_2, \cdots, x_n$ 的二次型也必须解决这一问题：即不论用何种非异坐标变换化 $\sum_{i,j=1}^{n} a_{ij}x_ix_j$ 为不同的标准形，其正项个数与负项个数都不会改变。为了解决这一问题，先证：

引理 1 设 β 与 δ 分别是 s 维与 t 维实的列向量，$k_i>0$，$i=1, 2, \cdots, s$；$l_j>0$，$j=1, 2, \cdots, t$，且满足

$$\begin{aligned}&\beta'[k_1, k_2, \cdots, k_s]\beta\\ &= -\delta'[l_1, l_2, \cdots, l_t]\delta,\end{aligned} \tag{1}$$

则 $\beta=0$。

【证明】 记 $\beta=(b_1, b_2, \cdots, b_s)$，$\delta=(d_1, d_2, \cdots, d_t)$，则(1)式化为

$$\sum_{i=1}^{s} k_ib_i^2+\sum_{j=1}^{t} l_jd_j^2=0,$$

由假设 $k_i>0$，$l_j>0$，而 b_i，d_j 都是实数，所以 $b_i=0$，$i=1, 2, \cdots, s$，即 $\beta=0$。 证毕。

定理 2.1 设 A 是秩为 r 的 n 阶实对称阵，如果 n 阶阵 P 与 Q 使

$$P'AP=[k_1, \cdots, k_p, -k_{p+1}, \cdots, -k_r, 0, \cdots, 0],$$
$$k_i>0,\ i=1, 2, \cdots, r$$
$$Q'AQ=[l_1, \cdots, l_q, -l_{q+1}, \cdots, -l_r, 0, \cdots, 0],$$
$$l_j>0,\ j=1, 2, \cdots, r$$

则必 $p=q$。

【证明】 由假设可得

$$\begin{aligned}&(P^{-1})'[k_1, \cdots, k_p, -k_{p+1}, \cdots, -k_r, 0, \cdots, 0]P^{-1}\\ &=(Q^{-1})'[l_1, \cdots, l_q, -l_{q+1}, \cdots, -l_r, 0, \cdots, 0]Q^{-1},\end{aligned} \tag{2}$$

把 P^{-1}、Q^{-1} 如下分块：

$$P^{-1} = \begin{matrix} & n \\ {}^{p} & \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \\ {}_{n-p} & \end{matrix}, \quad Q^{-1} = \begin{matrix} & n \\ {}^{q} & \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \\ {}_{n-q} & \end{matrix},$$

则(2)式化为

$$\begin{aligned} &P_1'[k_1, \cdots, k_p]P_1 + P_2'[-k_{p+1}, \cdots, -k_r, 0, \cdots, 0]P_2 \\ &= Q_1'[l_1, \cdots, l_q]Q_1 + Q_2'[-l_{q+1}, \cdots, -l_r, 0, \cdots, 0]Q_2. \end{aligned} \tag{3}$$

今先证：$p \leqslant q$。如果 $p > q$，则

$$(n-p)+q = n-(p-q) < n,$$

故齐次线性方程组：

$$\begin{matrix} & n \\ {}^{q} & \begin{pmatrix} Q_1 \\ P_2 \end{pmatrix} x = 0 \\ {}_{n-p} & \end{matrix}$$

必有非零解：$x = \delta \neq 0$，故

$$Q_1\delta = 0, \tag{4}$$

$$P_2\delta = 0, \tag{5}$$

在(3)式两边右乘 δ，左乘 δ'，並由(4)式与(5)式即得

$$\begin{aligned} &(P_1\delta)'[k_1, \cdots, k_p](P_1\delta) \\ &\quad = -(Q_2\delta)'[l_{q+1}, \cdots, l_r, 0, \cdots, 0](Q_2\delta). \end{aligned}$$

由于 $k_i > 0$, $l_j > 0$，故由引理 1，$P_1\delta = 0$，由此式及(5)式即得

$$\delta = P(P^{-1}\delta) = P\begin{pmatrix} P_1\delta \\ P_2\delta \end{pmatrix} = P\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0,$$

此为矛盾，故 $p \leqslant q$。同法可证：$q \leqslant p$。所以 $p = q$。　　　　证毕。

把定理 2.1 写成二次型的结论，就是下面的：

定理 2.1 （Sylvester 惯性定律）秩为 r 的实二次型 $x'Ax$ 的各个标准形的正项个数，因而负项个数相同。也即，若两个非异坐标变换 $x = Py$, $x = Qz$，使

$$x'Ax = \sum_{i=1}^{p} k_i y_i^2 - \sum_{i=p+1}^{r} k_i y_i^2 = \sum_{j=1}^{q} l_j z_j^2 - \sum_{j=q+1}^{r} l_j z_j^2,$$

而 $k_i > 0$, $l_j > 0$; $i, j = 1, 2, \cdots, r$，则必 $p = q$。

称 p 为实二次型 $x'Ax$ 的**正惯性指数**，$r-p$ 为 $x'Ax$ 的**负惯性指数**，$s=p-(r-p)=2p-r$ 为 $x'Ax$ 的**符号差**（即 $x'Ax$ 的标准形的正项个数减去负项个数）。由定理 2.1′ 可知，由于 r 与 p 由 $x'Ax$ 所唯一确定，故 s 也是由 $x'Ax$ 所唯一确定。

注意：定理 2.1′ 所说的二次型是实的，即不仅 A 是实方阵，且 x 也是实向量，非异坐标变换 $x=Qy$ 也是实的，即 Q 是非异实方阵，y 是实向量。但如果把实二次型看作复数域上的二次型，並且取 Q 为非异复方阵，则定理 2.1′ 不再成立。因为，如果 $x=Qy$ 使

$$x'Ax=\sum_{i=1}^{p}k_iy_i^2-\sum_{i=p+1}^{r}k_iy_i^2,\ k_i>0,\ i=1,\ 2,\ \cdots,\ r,$$

则令

$$P=[k_1^{-\frac{1}{2}},\ \cdots,\ k_p^{-\frac{1}{2}},\ (-k_{p+1})^{-\frac{1}{2}},\ \cdots,\ (-k_r)^{-\frac{1}{2}},\ 1,\ \cdots,1],$$
$$y=Pz,$$

于是 $x=(QP)z$ 是复的非异坐标变换，且使得

$$x'Ax=z_1^2+\cdots+z_p^2+z_{p+1}^2+\cdots+z_r^2,$$

所以惯性定律不再成立。

推论 1 设秩为 r 的实二次型 $x'Ax$ 的正惯性指数是 p，则必可找到非异坐标变换 $x=Ry$，使

$$x'Ax=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2. \tag{6}$$

【证明】 由假设，必可找到非异阵 Q，使

$$Q'AQ=[d_1,\ \cdots,\ d_p,\ -d_{p+1},\ \cdots,\ -d_r,\ 0,\ \cdots,\ 0],$$
$$d_i>0,\ i=1,\ 2,\ \cdots,\ r.$$

作方阵

$$R=Q[d_1^{-\frac{1}{2}},d_2^{-\frac{1}{2}},\ \cdots,\ d_r^{-\frac{1}{2}},\ 1,\ \cdots,\ 1],$$

则显然有

$$\begin{aligned}R'AR&=[1,\ \cdots,\ 1,\ -1,\ \cdots,\ -1,\ 0,\ \cdots,\ 0]\\&=[I_p,\ -I_{r-p},\ 0],\end{aligned} \tag{7}$$

故 $x=Ry$ 使成立(6)式。 证毕。

称(6)式右边的二次型为 $x'Ax$ 的**规范形**，而称 $[I_p,\ -I_{r-p},0]$ 为 A 的规范形。

推论 2 设实二次型 $x'Ax$ 与 $y'By$ 有相同的秩 r 与正惯性指数 p，则必存在非异坐标变换 $x=Qy$，使

$$x'Ax=y'By. \tag{8}$$

【证明】 由假设，存在非异阵 R 与 S，使

$$R'AR=[I_p,\ -I_{r-p},\ 0],$$

$$S'BS=[I_p,\ -I_{r-p},\ 0].$$

令 $Q=RS^{-1}$，则 $Q'AQ=B$，于是 $x=Qy$ 使(8)式成立。 证毕。

称 $x'Ax$ 与 $y'By$ 可以**互化**，或称它们是**等价的**，如果存在非异坐标变换 $x=Qy$ 使成立(8)式。

推论 2 以及惯性定律表明：两个实二次型等价的充要条件是，它们有相同的秩与正惯性指数。

二、(实)二次型的分类

如果把所有等价的实二次型归入同一类（也即它们的系数矩阵相互合同的归入同一类，即合同类），则可知同一类中的任意两个二次型必定等价（可以互化），而不同类中的任意两个二次型决不等价，故由上段的讨论可知，二次型的秩 r 与正惯性指数 p 完全决定了二次型的分类。于是可把实二次型分成三大类：

(i) $p=r\leqslant n$，即 $x=Qy$，使

$$x'Ax=y_1^2+y_2^2+\cdots+y_r^2,\ r\leqslant n, \tag{9}$$

或者，

$$Q'AQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}=[I_r,\ 0],\ r\leqslant n. \tag{10}$$

此时称 $x'Ax$ 为**半正定(或非负定)二次型**，而称 A 为**半正定(非负定)阵**。

当 $p=r=n$，即 $x=Qy$，使

$$x'Ax=y_1^2+y_2^2+\cdots+y_n^2. \tag{11}$$

或者

$$Q'AQ=I_n. \tag{12}$$

此时称 $x'Ax$ 为**正定二次型**，而称 A 为**正定阵**。

(ii) $p=0, r\leqslant n$，即 $x=Qy$，使

$$x'Ax=-y_1^2-y_2^2-\cdots-y_r^2,\ r\leqslant n. \tag{13}$$

或者

$$Q'AQ=[-I_r, 0],\ r\leqslant n. \tag{14}$$

此时称 $x'Ax$ 为**半负定二次型**，而称 A 为**半负定阵**。

当 $r=n$ 时，称 $x'Ax$ 为**负定二次型**，而称 A 为**负定阵**。

(iii) $0<p<r\leqslant n$，即 $x'Ax$ 的规范形是

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2(=x'Ax),\ r\leqslant n. \tag{15}$$

此时称 $x'Ax$ 为**不定二次型**，而称 A 为**不定阵**。

容易看出，如果 $x'Ax$ 是负定二次型，则 $x'(-A)x$ 是正定二次型，如果 $x'Ax$ 是半负定二次型，则 $x'(-A)x$ 是半正定二次型，所以只要讨论正定二次型、半正定二次型、不定二次型这三类二次型就可以了。而本章以下只讨论应用最多的前两种类型的二次型。

习　　题

1. 把下列各个二次型化成标准形，並进行分类：

(i) $4x_1x_2-2x_2x_3+6x_1x_3$；

(ii) $4x_1^2-2x_1x_3+6x_2^2-8x_3^2$；

(iii) $\sum\limits_{i=1}^{n}x_i^2+\sum\limits_{i=1}^{n-1}x_ix_{i+1}$；

(iv) $n\sum\limits_{i=1}^{n}x_i^2-\left(\sum\limits_{i=1}^{n}x_i\right)^2$.

2. 求证：非零的实二次型可分解为两个实线性型的乘积的充要条件是，$x'Ax$ 的秩等于 2 以及符号差等于零；或者，$x'Ax$ 的秩等于 1.

3. 设 a, b 是实数，求实二次型

$$a\sum_{i=1}^{n} x_i^2+2b\sum_{1\leqslant i<j\leqslant n}x_ix_j$$

的秩与符号差。

4. 设 n 个变量 $x_1, x_2, \cdots, x_n$ 的实二次型是

$$x'Ax=l_1^2+\cdots+l_p^2-l_{p+1}^2-\cdots-l_r^2(p\leqslant r\leqslant n),$$

其中 $l_i=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n$, $i=1, 2, \cdots, r$, 求证：$x'Ax$ 的正惯性指

数不大于 p，而负惯性指数不大于 $r-p$。

5. 设H是 n 阶实镜象阵，求实二次型 $x'Hx$ 的正惯性指数与符号差。

§3 正定二次型与正定阵

正定二次型在二次型理论中，正定阵在矩阵论中占有特殊重要的位置，本节将详细讨论它们的一些基本性质。

定理 3.1 二次型 $x'Ax$ 是正定(二次型)的充要条件是，对任何 n 维实的非零列向量 x，必有 $x'Ax>0$。

【证明】 必要性。由假设 $x'Ax$ 是正定二次型，故存在实的非异坐标变换 $x=Qy$，$y=(y_1, y_2, \cdots, y_n)'$，使

$$x'Ax=y_1^2+y_2^2+\cdots+y_n^2. \tag{1}$$

对 $x\neq 0$，因 Q 非异，故 $y\neq 0$，于是由(1)可知，$x'Ax>0$。

充分性。设 $x'Ax$ 的秩与正惯性指数分别为 r 与 p，先证 $r=p$，如 $p<r$，则由惯性定律，存在非异坐标变换：$x=Qy$，使

$$x'Ax=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2. \tag{2}$$

由假设，对任何 $x\neq 0$，$x'Ax>0$，但对列向量

$$x=Q(0,\cdots,0,1,0,\cdots,0)'\neq 0$$

(因 Q 是非异阵，1 是 x 的第 $p+1$ 个分量). 却有

$$x'Ax=-1<0,$$

这与假设相矛盾。故 $r=p$。再证 $r=n$。如果 $r<n$，则(2) 式应化为

$$x'Ax=y_1^2+y_2^2+\cdots+y_r^2,\ r<n. \tag{3}$$

于是取

$$x=Q(0,\cdots,0,1)'\neq 0.$$

由(3)即得 $x'Ax=0$，又与假设相矛盾，故 $r=n=p$，即 $x'Ax$ 是正定二次型。 证毕。

例 1 设 A 与 B 都是正定阵，则 $A+B$ 也是正定阵。

【证明】 因 A 与 B 均正定，故由定理 3.1 的必要条件可知，对任何 $x\neq 0$，必有 $x'Ax>0$，$x'Bx>0$，所以

$$x'(A+B)x = x'Ax + x'Bx > 0.$$

显然 $A+B$ 也是实对称阵，于是由定理3.1的充分条件可知，$A+B$ 是正定阵。 证毕。

例 2 设 A 是正定阵，P 是非异实方阵，则 $P'AP$ 也是正定阵。

【证明】 因为 A 是实对称阵，故 $P'AP$ 显然也是实对称阵。又对任何实的列向量 $x \neq 0$，由于 $Px \neq 0$（因 P 是非异阵），故

$$x'(P'AP)x = (Px)'A(Px) > 0,$$

即 $P'AP$ 是正定阵。 证毕。

定理 3.2 n 阶实对称阵 A 是正定阵的充要条件是，A 的特征值全大于零。

【证明】 由定理 1.1，存在正交坐标变换 $x = Qy$，使

$$x'Ax = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2. \tag{4}$$

充分性。对任何 $x \neq 0$，则 $y = Q'x \neq 0$，而由假设 $\lambda_i > 0$，$i = 1, 2, \cdots, n$，故由（4）式即得 $x'Ax > 0$，于是由定理 3.1，$x'Ax$ 是正定二次型，即 A 是正定阵。

必要性。若 A 是正定阵，则由定理 3.1，对任何 $x \neq 0$，恒有 $x'Ax > 0$，于是只要有一个 λ_i，例如 $\lambda_1 \leqslant 0$，则可取非零向量 $x = Q(1, 0, \cdots, 0)'$，而 $x'Ax = \lambda_1 \leqslant 0$，这是矛盾的，所以 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 全大于零。 证毕。

例 1、例 2 与定理 3.2 的证明说明，在论证对称阵是否正定阵时，可以用它相应的二次型来推导，而不必拘泥于用纯矩阵的方法，以用何种方法最简捷为准。

定理 3.3 设 A 是 n 阶正定阵，则

(i) $A = B'B$，而 B 是非异阵， (5)

(ii) $A = R'R$，而 R 是主对角元全大于零的上三角阵， (6)

称(6)式为 A 的 **Cholesky 分解**。

(iii) A 的所有主子式全大于零。

［证明］ 因为 A 是正定阵，故存在非异阵 Q，使

$$Q'AQ = I_n,$$

于是

$$A=(Q^{-1})'Q^{-1}=B'B,$$

其中 $B=Q^{-1}$ 是非异阵。这证明了(i)。

再证(ii)。由(i)作 B 的 QR 分解：$B=QR$，其中 Q 是正交阵，R 是主对角元全大于零的上三角阵，故

$$A=B'B=R'Q'QR=R'R.$$

今证(iii)。由(ii)及 Cauchy-Binet 公式可知，对 A 的任何 k 阶主子式，$1\leqslant k\leqslant n$，恒有下式：

$$\begin{aligned}\left|A\begin{pmatrix}i_1i_2\cdots i_k\\ i_1i_2\cdots i_k\end{pmatrix}\right|&=\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\left|R'\begin{pmatrix}i_1i_2\cdots i_k\\ j_1j_2\cdots j_k\end{pmatrix}\right|\left|R\begin{pmatrix}j_1j_2\cdots j_k\\ i_1i_2\cdots i_k\end{pmatrix}\right|\\&=\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\left|R\begin{pmatrix}j_1j_2\cdots j_k\\ i_1i_2\cdots i_k\end{pmatrix}\right|^2\\&=\left|R\begin{pmatrix}i_1i_2\cdots i_k\\ i_1i_2\cdots i_k\end{pmatrix}\right|^2+\zeta(\zeta\geqslant 0).\end{aligned}\tag{7}$$

因为 R 是非异上三角阵，故 $R\begin{pmatrix}i_1i_2\cdots i_k\\ i_1i_2\cdots i_k\end{pmatrix}$ 也是非异上三角阵，所以由(7)式可知

$$\left|A\begin{pmatrix}i_1i_2\cdots i_k\\ i_1i_2\cdots i_k\end{pmatrix}\right|>0,\ 1\leqslant i_1<i_2<\cdots<i_k\leqslant n.$$ 证毕。

推论 1 正定阵的主对角元必全大于零。

这由定理 3.3 的(iii)可知是显然的。推论 1 主要用于判定某些实对称阵不是正定阵。例如，主对角元有非正数的对称阵必不是正定阵。

推论 2 设 A 是正定阵，则 A 的元素的绝对值最大者必是主对角元。

【证明】 设 $|a_{ij}|$ 最大，$i\neq j$，则 $|a_{ij}|^2\geqslant a_{ii}a_{jj}$。而由定理 3.3 的(iii)可知

$$a_{ii}a_{jj}-|a_{ij}|^2=\begin{vmatrix}a_{ii} & a_{ij}\\ a_{ij} & a_{jj}\end{vmatrix}>0,$$

故得 $|a_{ij}|^2<a_{ii}\cdot a_{jj}$，此为矛盾，故 A 中元素的绝对值最大者必是

主对角元。 证毕。

推论 2 主要也是用于判定某些实对称阵不是正定阵，因为只要有一个非主对元的绝对值不小于主对角元的最大者，则这个实对称阵必不是正定阵。

定理 3.4 只要 n 阶实对称阵 A 满足下列条件中的任意一个：

(i) $A=B'B$，而 B 是 $m\times n$ 列满秩阵；

(ii) A 的所有 k 阶主子式之和大于零，$k=1, 2, \cdots, m$，则 A 必是正定阵。

【证明】 因(i)成立，故

$$x'Ax=x'B'Bx=(Bx)'(Bx), \tag{8}$$

当 $x\neq 0$ 时，必有 $Bx\neq 0$。因若 $Bx=0=B\cdot 0$，则因 B 是列满秩阵，可从矩阵等式左侧消去，故得 $x=0$，此为矛盾。由 (8) 式以及 $Bx\neq 0$ 即知 $x'Ax>0$，由定理 3.1，A 是正定阵。

再证(ii)。设 A 的特征多项式是

$$|\lambda I_n-A|=\lambda^n-a_1\lambda^{n-1}+\cdots+(-1)^{n-1}a_{n-1}\lambda+(-1)^n a_n, \tag{9}$$

则由第五章定理，a_k 是 A 的所有 k 阶主子式之和，故由假设，$a_k>0$, $k=1, 2, \cdots, n$。于是 A 的特征值全不为零（否则会产生 $a_n=|A|=0$ 的矛盾）；又 A 的任一特征值不能小于零。因如不然：$\lambda=-c$, $c>0$，则由(9)式，

$$0=|-cI_n-A|=(-1)^n(c^n+a_1c^{n-1}+\cdots+a_{n-1}c+a_n),$$

但上式右端不等于零(因 $c>0$, $a_k>0$, $k=1, 2, \cdots, n$)此为矛盾。所以 A 的特征值全大于零，由定理 3.2，A 是正定阵。 证毕。

定理 3.5 如果 n 阶实对称阵 A 的 n 个顺序主子式全大于零，则 A 必是正定阵。

【证明】 对 n 用归纳法，当 $n=1$ 时，定理显然正确。今把 A 如下分块：

$$A=\begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix},$$

其中 A_{n-1} 是 $n-1$ 阶顺序主子阵，故由 §1 引理 1，存在非异阵 Q，

使

$$Q'AQ = [A_{n-1},\ a_{nn} - \alpha' A_{n-1}^{-1}\alpha].$$

因为 A_{n-1} 的 $n-1$ 个顺序主子式就是 A 的前 $n-1$ 个顺序主子式，它们全大于零，故由归纳法假设，A_{n-1} 是正定阵，于是存在 $n-1$ 阶非异阵 Q_1，使 $Q_1'A_{n-1}Q_1 = I_{n-1}$，记 $R = Q[Q_1,\ 1]$，则

$$\begin{aligned} R'AR &= [Q_1,\ 1]'[A_{n-1},\ a_{nn} - \alpha' A_{n-1}^{-1}\alpha][Q_1,\ 1] \\ &= [I_{n-1},\ k], \end{aligned} \tag{10}$$

其中 $k = a_{nn} - \alpha' A_{n-1}^{-1}\alpha$，由行列式的第一降阶定理，$k = \dfrac{|A|}{|A_{n-1}|}$，由假设 $|A| > 0$，$|A_{n-1}| > 0$，故 $k > 0$，于是记 $S = R[I_{n-1},\ k^{-1/2}]$，则 S 是非异阵，且（10）式可改写为：$S'AS = I_n$。所以 A 是正定阵。 证毕。

例 3 正定阵 A 的任何主子阵必是正定阵。

【证明】 设 $A\begin{pmatrix} i_1 i_2 \cdots i_k \\ i_1 i_2 \cdots i_k \end{pmatrix}$ 是 n 阶正定阵 A 的主子阵，则它也是实对称阵，且它的 k 个顺序主子式也是 A 的 k 个主子式，而由定理 3.3 的 (iii)，它们都大于零，故由定理 3.5，$A\begin{pmatrix} i_1 i_2 \cdots i_k \\ i_1 i_2 \cdots i_k \end{pmatrix}$ 是正定阵。 证毕。

例 4 设 $a_i > 0$，且 a_i 全不相同，$i = 1, 2, \cdots, n$，则 n 阶阵：$A = \left| \dfrac{1}{a_i + a_j} \right|_{n\times n}$ 必是正定阵。

【证明】 A 显然是对称阵。又 A 的任一 k 阶顺序主子式必是 $\left| \dfrac{1}{a_i + a_j} \right|_{k\times k}$，$1 \leqslant k \leqslant n$，这个行列式是 Cauchy 行列式，它应是：

$$\left| \frac{1}{a_i + a_j} \right|_{k\times k} = \frac{\prod\limits_{1\leqslant i<j\leqslant k} (a_i - a_j)^2}{\prod\limits_{i,j=1}^{n} (a_i + a_j)} > 0$$

（参阅第二章选做题第 8 题）。所以由定理3.5，A 必是正定阵。证毕。

注意：定理 3.3、定理 3.4 与定理 3.5 的任何一个都可作为判断实对称阵是否正定阵的充分必要条件，这里把它分开，是为了要

在较弱条件下，得到较强的结果。

将在选做题 2 中给出判断对称阵是正定阵的另一充分条件。

习　题

1. 设 A 是 n 阶正定阵，求证下列各方阵都是正定阵。

(i) $\mathrm{adj}A$（$\mathrm{adj}A$ 是 A 的伴随阵）；

(ii) A^k（k 是整数）；

(iii) $lA(l>0)$；

(iv) $B'AB$（B 是 $n\times m$ 行满秩阵）；

(v) $(a_{ij}b_ib_j)_{n\times n}$, $b_i>0$, $i=1, 2, \cdots, n$, 其中 $A=(a_{ij})_{n\times n}$。

2. 设 A 是 n 阶实对称阵，求证：

(i) 必存在一个充分小的正数 ϵ，使 $I_n+\epsilon A$ 是正定阵。

(ii) 必存在一个充分大的正数 N，使 NI_n+A 是正定阵。

3. 设 A 是秩为 r 的 n 阶实方阵，$A=P\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}Q$，其中 P 与 Q 都是 n 阶非异阵，试应用 $P'P$ 与 $Q'Q$ 的正定性分别证明：

$$\mathrm{r}(A'A)=\mathrm{r}(A),\ \mathrm{r}(AA')=\mathrm{r}(A),$$

此处 $\mathrm{r}(A)$ 表示 A 的秩。

4. 设 B 是 $n\times m$ 阵，A 是 n 阶正定阵，求证：

$$\mathrm{r}(B'AB)=\mathrm{r}(B).$$

5. 求证：n 阶实对称阵 A 是正定阵的充分必要条件是，A 的 n 个顺序主子式的代数余子式全大于零。

6. 求证：实对称阵 A 是正定阵的充要条件是，$A=UU'$，而 U 是非异上三角阵。

7. 设 A 是 n 阶实对称阵，证明以下四条是等价的：

(i) A 是负定阵；

(ii) $(-1)^i\left|A\begin{pmatrix}12\cdots i\\ 12\cdots i\end{pmatrix}\right|>0$, $i=1, 2, \cdots, n$；

(iii) $A=-GG'$，而 G 是非异下三角阵；

(iv) $A=UU'$，而 U 是非异上三角阵。

8. 求证：主对角元全正的严格对角占优的实对称阵必是正定阵。（注：某些未必是严格对角占优，但主对角元之和占优的实对称阵也可能是正定阵——见本章选做题第 2 题。）

§4 半正定二次型、正定二次型(正定阵)的应用

一、二次型是半正定的充分或必要条件

由于本段前两个定理的证明与§3前两个定理的证法相同，故只列出结论而略去其证明。

定理 4.1 实二次型是半正定的充要条件是，对任何 n 维实的列向量 x，恒有 $x'Ax\geqslant 0$。

定理 4.2 实对称阵 A 是半正定阵的充要条件是，A 的任何特征值都是非负数。

定理 4.3 实对称阵 A 是秩为 r 的半正定阵的充要条件是，存在 $r\times n$ 行满秩阵 B，使

$$A=B'B. \tag{1}$$

【证明】 如存在 $r\times n$ 行满秩阵 B 使成立(1)式，则对任何 n 维实的列向量 x，恒有

$$x'Ax=(Bx)'(Bx)\geqslant 0,$$

故由定理 4.1，A 是半正定阵，且

$$\mathrm{r}(A)=\mathrm{r}(B'B)=\mathrm{r}(B)=r.$$

反之，如果 A 是秩为 r 的半正定阵，则存在非异阵 P，使得

$$A=P'^{-1}\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}P^{-1}. \tag{2}$$

把 P^{-1} 如下分块：

$$P^{-1}=\begin{matrix} \\ r\\ \\ \end{matrix}\overset{n}{\begin{pmatrix}B\\ C\end{pmatrix}},$$

则 B 是 $r\times n$ 行满秩阵，且(2)式化为

$$A=(B',C')\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}B\\ C\end{pmatrix}=B'B.$$ 证毕。

如果把定理 4.3 中的 $n\times r$ 列满秩阵 B' 按它的列分块：$B=$

$(\beta_1, \beta_2, \cdots, \beta_r)$，则 $\beta_1, \beta_2, \cdots, \beta_r$ 线性无关，且由(1)式显然可得

推论 秩为 r 的半正定阵必可分解为 r 个秩为 1 的半正定阵之和，即：

$$A = \beta_1\beta_1' + \beta_2\beta_2' + \cdots + \beta_r\beta_r', \tag{3}$$

其中 $\beta_1, \beta_2, \cdots, \beta_r$ 是线性无关的 n 维实的列向量。

定理 4.4 如果 n 阶实对称阵 A 的所有 k 阶主子式之和是非负数，$k=1, 2, \cdots, n$，则 A 是半正定阵。

由于证法与定理 3.4 的(ii)的证明相同，故从略。

注意：不能把定理 3.5 的条件相应地搬过来。也就是，即使实对称阵 A 的所有顺序主子式全是非负数，A 也未必是半正定阵。例如，

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

的所有顺序主子式 Δ_1, Δ_2 全是非负数：$\Delta_1 = \Delta_2 = 0$，但对向量 $x = (0, b)'$，$b \neq 0$，$x'Ax = -b^2 < 0$，故由定理 4.1，A 必不是半正定阵。

定理 4.5 n 阶实对称阵 A 是半正定阵的充要条件是，A 的所有主子式全是非负数。

【证明】 如果 A 是秩为 r 的半正定阵，则由定理 4.3，$A = B'B$，而 B 是 $r \times n$ 行满秩阵。于是与定理 3.3 的 (iii) 的证法相仿，可证得 A 的任一主子式必是非负数。

反之，如果实对称阵 A 的所有主子式都是非负数，则 A 的所有 k 阶主子式之和也是非负数，$k=1, 2, \cdots, n$。故由定理 4.4，A 是半正定阵。 证毕。

不难看出，定理 4.5 的充分条件比定理 4.4 的条件要强些，故运用定理 4.4 的条件可能更好些(灵活性大些)。

二、Schur 定理及其应用

设 $A = (a_{ij})_{n\times n}$，$B = (b_{ij})_{n\times n}$，记 $A*B = (a_{ij}b_{ij})_{n\times n}$，称 $A*B$ 为 A 与 B 的 **Hadamard 乘积** (或 **Schur 乘积**)(参阅第三章选做题

第 15 题)。

定理 4.6(Schur) 设 A 与 B 都是 n 阶半正定阵，则 $A*B$ 也是半正定阵。当 A 与 B 都是正定阵时，则 $A*B$ 也是正定阵。

【证明】 设 B 的秩为 r，则由定理 4.3，存在 $r\times n$ 行满秩阵 G，使 $B=G'G$。记 $G'=(g_{ij})_{n\times r}$，$B=(b_{ij})_{n\times n}$，则可得：

$$b_{ij}=\sum_{k=1}^{r}g_{ik}g_{jk},\ i,\ j=1,\ 2,\ \cdots,\ n. \tag{4}$$

设 $x=(x_1,\ x_2,\ \cdots,\ x_n)'$，则二次型

$$\begin{aligned}x'(A*B)x&=\sum_{i,j=1}^{n}x_ia_{ij}b_{ij}x_j\\&=\sum_{i,j=1}^{n}\sum_{k=1}^{r}x_ia_{ij}g_{ik}g_{jk}x_j\\&=\sum_{k=1}^{r}\sum_{i,j=1}^{n}a_{ij}(x_ig_{ik})(x_jg_{jk}),\end{aligned} \tag{5}$$

记

$$y_i^{(k)}=x_ig_{ik},\ i=1,\ 2,\ \cdots,\ n;\quad k=1,\ 2,\ \cdots,\ r, \tag{6}$$

则(5)式可化为

$$\begin{aligned}x'(A*B)x&=\sum_{k=1}^{r}(y_1^{(k)},\ \cdots,\ y_n^{(k)})A(y_1^{(k)},\ \cdots,\ y_n^{(k)})'\\&=\sum_{k=1}^{r}Y_k'AY_k,\end{aligned} \tag{7}$$

其中

$$Y_k=(y_1^{(k)},\ \cdots,\ y_n^{(k)})',\ k=1,\ 2,\ \cdots,\ r. \tag{8}$$

由于 A 是半正定阵，故由定理 4.1，$Y_k'AY_k\geqslant 0$，$k=1,\ 2,\ \cdots,\ r$，所以由(7)式可知，$x'(A*B)x\geqslant 0$，$A*B$ 显然也是实对称阵，故由定理 4.1，$A*B$ 是半正定阵。

当 A 与 B 都是正定阵时，由 B 的正定性可知，$r=n$（定理 3.3 的(i)），故(7)式应改写成：

$$x'(A*B)x=\sum_{k=1}^{n}Y_k'AY_k. \tag{9}$$

又由(6)式与(8)式可得

$$(\mathbf{Y}_1, \mathbf{Y}_2, \cdots, \mathbf{Y}_n) = \begin{pmatrix} x_1 g_{11} & x_1 g_{12} & \cdots & x_1 g_{1n} \\ x_2 g_{21} & x_2 g_{22} & \cdots & x_2 g_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_n g_{n1} & x_n g_{n2} & \cdots & x_n g_{nn} \end{pmatrix}$$

$$= [x_1, x_2, \cdots, x_n] G'. \qquad (10)$$

因为 G' 是非异阵，故由 (10)式显然可知，当 $x \neq 0$ 时，n 阶阵$(\mathbf{Y}_1, \mathbf{Y}_2, \cdots, \mathbf{Y}_n)$中至少有某一列，例如 $\mathbf{Y}_s$，不等于零，$1 \leqslant s \leqslant n$，又由假设 A 是正定阵，故 $\mathbf{Y}_s' A \mathbf{Y}_s > 0$，所以由(9) 式可知，$x'(A*B)x > 0$，于是由定理 3.1，$A*B$ 是正定阵。　　　　证毕。

推论(华罗庚)　设 $A = (a_{ij})_{n\times n}$ 是正定阵，则对任何正整数 k，n 阶阵：

$$(a_{ij}^k)_{n\times n} = \begin{pmatrix} a_{11}^k & a_{12}^k & \cdots & a_{1n}^k \\ a_{21}^k & a_{22}^k & \cdots & a_{2n}^k \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}^k & a_{n2}^k & \cdots & a_{nn}^k \end{pmatrix}$$

必是正定阵。

【证明】　因为

$$(a_{ij}^k)_{n\times n} = \overbrace{(a_{ij})_{n\times n} * (a_{ij})_{n\times n} * \cdots * (a_{ij})_{n\times n}}^{k\text{个}},$$

故由 Schur 定理及归纳法即得证。　　　　证毕。

例 1　因为 $\left(\frac{1}{i+j}\right)_{n\times n}$ 是正定阵(由 §3 例 4)，所以由上述推论，方阵

$$\left(\frac{1}{(i+j)^5}\right)_{n\times n} = \begin{pmatrix} \frac{1}{2^5} & \frac{1}{3^5} & \cdots & \frac{1}{(n+1)^5} \\ \frac{1}{3^5} & \frac{1}{4^5} & \cdots & \frac{1}{(n+2)^5} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{(n+1)^5} & \frac{1}{(n+2)^5} & \cdots & \frac{1}{(2n)^5} \end{pmatrix}$$

是正定阵。

最后举一个综合运用 Schur 定理的例子：

例 2 设 n 阶实方阵 A 的 n 个特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是全不相同的负实数，C 是负定阵，则必存在 n 阶正定阵 B，使 $AB+BA'=C$。

【证明】 由假设 $\lambda_i<0, i=1, 2, \cdots, n$，故 $\lambda_i+\lambda_j\neq 0, i, j=1, 2, \cdots, n$。因而存在 n 阶实对称阵 B，使得 $AB+BA'=C$（见第六章选做题第 20 题）。又由假设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 全不相同，故由第五章定理 2.1 的推论 2，必存在非异阵 P，使 $PAP^{-1}=[\lambda_1, \lambda_2, \cdots, \lambda_n]$，故 $AB+BA'=C$ 可化为：

$$[\lambda_1, \lambda_2, \cdots, \lambda_n](PBP')+(PBP')[\lambda_1, \lambda_2, \cdots, \lambda_n] = PCP'. \quad (11)$$

记

$$PBP'=B_1,\ PCP'=C_1;\ B_1=(b_{ij})_{n\times n},\ C_1=(c_{ij})_{n\times n},$$

则(11)式化为：

$$((\lambda_i+\lambda_j)b_{ij})_{n\times n}=(c_{ij})_{n\times n}.$$

故由上式可得

$$b_{ij}=\frac{c_{ij}}{\lambda_i+\lambda_j}=\frac{-c_{ij}}{(-\lambda_i)+(-\lambda_j)},\quad i, j=1, 2, \cdots, n. \quad (12)$$

又由假设，C是负定阵，故 C_1 也是负定阵，因而 $-C_1$是正定阵，又记

$$F=\left(\frac{1}{(-\lambda_i)+(-\lambda_j)}\right)_{n\times n},$$

则因 $-\lambda_i>0, i=1, 2, \cdots, n$，故由 §3 例 4 可知，$F$ 是正定阵。由于(12)式可写成：

$$B_1=(-C_1)*F,$$

所以由 Schur 定理，B_1 也是正定阵。因而 $B=P^{-1}B_1(P^{-1})'$ 也是正定阵。 证毕。

三、正定二次型与正定阵的应用

正定二次型与正定阵有很多应用，上节已有所述及，今继续叙述其一部分应用，先看正定二次型在数学分析中的应用。

例 3 设 A 是 n 阶正定阵，实向量 $\alpha=(c_1, c_2, \cdots, c_n)'$ 是二次函数：

$$f(x)=f(x_1, x_2, \cdots, x_n)=\sum_{i,j=1}^{n} a_{ij}x_ix_j-2\sum_{i=1}^{n} b_ix_i \tag{13}$$

的极值点，则 α 必是 $f(x)$的极小点，且 $\alpha=A^{-1}b$。再求出 $f(\alpha)$，此处 $b=(b_1, b_2, \cdots, b_n)'$。

【证明】 **证法一**。由假设 $\alpha=(c_1, c_2, \cdots, c_n)'$ 是 $f(x)$ 的极值点，故

$$\left.\frac{\partial f}{\partial x_i}\right|_{x_j=c_j}=0, \qquad i, j=1, 2, \cdots, n.$$

也即

$$2\left(\sum_{j=1}^{n} a_{ij}x_j, -b_i\right)\Bigg|_{x_j=c_j}=0, \qquad i, j=1, 2, \cdots, n.$$

上式即

$$A\alpha=b. \tag{14}$$

今考虑在 α 的近旁 $\alpha+y$ 函数 $f(x)$的变化情形，此处 $y=(y_1, \cdots, y_n)'\neq 0$, $|y_i|<\epsilon_i$, $i=1, 2, \cdots, n$, 因为(13)式可以改写为：

$$f(x)=x'Ax-2b'x$$

故

$$f(\alpha+y)-f(\alpha)=(\alpha+y)'A(\alpha+y)-2b'(\alpha+y)$$
$$+2b'\alpha-\alpha'A\alpha,$$

以(14)式代入上式，并经过简单计算，上式可化为

$$f(\alpha+y)-f(\alpha)=y'Ay.$$

由假设 A 是正定阵，$y\neq 0$，故 $y'Ay>0$，因此 $f(\alpha+y)-f(\alpha)>0$，这说明 α 是 $f(x)$的极小点。且由(14)显然可得，$\alpha=A^{-1}b$。而由(15)式，$f(x)$的极小值 $f(x)$应是

$$f(\alpha)=-b'\alpha=-b'A^{-1}b.$$

证法二 今用纯代数的方法来证明。由证法一的结论所暗示，可把 $f(x)$改写成：

$$f(x)=(x-A^{-1}b)'A(x-A^{-1}b)-b'A^{-1}b \tag{15}$$

由假设 A 是正定阵，故 $(x-A^{-1}b)'A(x-A^{-1}b)\geqslant 0$，而当且仅当

$x-A^{-1}b=0$ 时，等号才成立。故 $f(x)$ 当 $x=A^{-1}b$ 时取极小值，且 $f(A^{-1}b)=-b'A^{-1}b$。

虽然证法二要简捷得多，但(15)式的建立却非唾手可得的，而是在证法一的基础上才比较自然地建立起来。

下面继续把正定阵的理论用在行列式论上。

例 4 设 B 是非异实对称阵，C 是反对称实方阵，且 $BC=CB$，求证：$B+C$ 必是非异阵。

【证明】 对任何非零列向量 x，由于

$$x'(B+C)(B+C)'x=x'BB'x+x'CC'x+x'(CB'+BC')x. \tag{16}$$

再由假设，B 是对称阵，C 是反对称阵，以及 $BC=CB$，故

$$CB'+BC'=CB-BC=0,$$

又因 B 是非异阵，故 BB' 是正定阵，即 $x'BB'x>0$。又 CC' 是半正定阵，即 $x'CC'x\geqslant 0$，故(16)式应是

$$x'(B+C)(B+C)'x>0,$$

所以 $(B+C)(B+C)'$ 是正定阵，于是 $|B+C|^2>0$，因此 $B+C$ 是非异阵。 证毕。

例 5 设

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ A_{s1} & A_{s2} & \cdots & A_{ss} \end{pmatrix} \tag{17}$$

是正定阵，A_{ii} 均是方阵(未必同阶)，$i=1, 2, \cdots, s$，则必有：

$$|A|\leqslant|A_{11}||A_{22}|\cdots|A_{ss}|, \tag{18}$$

而等号成立的充要条件是：$A_{ij}=0$，$i\neq j$，$i, j=1, 2, \cdots, s$。

【证明】 对 s 用归纳法。当 $s=1$，结论显然正确。今把分块阵(17)写成：

$$A=\begin{pmatrix} A_{11} & B \\ B' & D \end{pmatrix},$$

其中

$$D=\begin{pmatrix} A_{22} & \cdots & A_{2s} \\ \cdots & \cdots & \cdots \\ A_{s2} & \cdots & A_{ss} \end{pmatrix},\ B=(A_{12},A_{13},\cdots,A_{1s}).$$

因 A_{11} 是 A 的顺序主子阵，故它是正定阵(自然是非异阵)，由 § 1 的引理 1，存在非异阵 P，使

$$P'AP=[A_{11},\ D-B'A_{11}^{-1}B].$$

因为 $P'AP$ 也是正定阵，而 $D-B'A_{11}^{-1}B$ 是它的主子阵，故由 §3 例 3 可知，$D-B'A_{11}^{-1}B$ 也是正定阵。又 A_{11}^{-1} 是正定阵，故当 $B\neq 0$ 时，$B'A_{11}^{-1}B$ 是半正定阵。于是当 $B\neq 0$ 时，

$$\begin{aligned} |D| &= |(D-B'A_{11}^{-1}B)+B'A_{11}^{-1}B| \\ &> |D-B'A_{11}^{-1}B|+|B'A_{11}^{-1}B| \end{aligned} \tag{19}$$

(因为正定阵加非零半正定阵的行列式大于这两个方阵各自的行列式之和。作为习题，见本节习题第 7 题)。(19)式还可写成：

$$|D|>|D-B'A_{11}^{-1}B|, \tag{20}$$

于是对任意的 B，由行列式的第一降阶定理以及(20)式即得

$$|A|=|A_{11}|\,|D-B'A_{11}^{-1}B|\leqslant|A_{11}|\,|D|. \tag{21}$$

而等式成立的充要条件是 $B=(A_{12},\cdots,A_{1s})=O$。又由归纳法假设，

$$|D|\leqslant|A_{22}|\,|A_{33}|\cdots|A_{ss}|. \tag{22}$$

而等号成立的充要条件是 $A_{ij}=0,\ i\neq j,\ i,\ j=2,3,\cdots,s$。所以由(21)式与(22)式即得本例要证的结论。　　证毕。

习　题

1. 设 A 与 B 是半正定阵，求证：$A+B$ 也是半正定阵。如果 A 与 B 中有一个是正定阵，则 $A+B$ 也是正定阵。

2. 设 A 是 n 阶实对称阵，且 $|A|<0$，则必存在 n 维实的列向量 $\alpha\neq 0$，使 $\alpha'A\alpha<0$。

3. 求证：实二次型：$n\sum\limits_{i=1}^{n}x_i^2-\left(\sum\limits_{i=1}^{n}x_i\right)^2$ 是半正定二次型。

4. 设 $a_i>0$，且 a_i 全不相同，$i=1,2,\cdots,n$，求证：方阵

$$\left(\frac{1}{(i+j)(a_i+a_j)^2}\right)_{n\times n}$$

是正定阵。

5. 设A是非零的半正定阵，求证：$\mathrm{tr}A>0$。

6. 设A是秩为r的n阶实对称阵，求证：A是幂等阵的充要条件是：$A=B(B'B)^{-1}B'$，而B是$n\times r$列满秩阵。

7. 设A是正定阵，B是半正定阵，且$B\neq 0$，求证：$|A+B|>|A|+|B|$。(提示：把A合同于单位阵：$P'AP=I$，然后证明$|I+P'BP|$大于

$$|I|+|P'BP|.)$$

8. 设$A=(a_{ij})_{n\times n}$是正定阵，求证：

$$|A|\leqslant a_{11}a_{22}\cdots a_{nn}.$$

而等号成立的充要条件是，$a_{ij}=0$，$i\neq j$，$i, j=1, 2, \cdots, n$。

9. 估计下面行列式的界限：

(i) (Hadamard)设$A=(a_{ij})_{n\times n}$是任意实方阵，求证：

$$|A|\leqslant\left(\prod_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2\right)^{1/2}.$$

(ii) 求证：

$$\begin{vmatrix} 3 & 3 & 2 & 2 \\ 1 & 4 & 3 & 2 \\ 1 & -2 & 4 & 3 \\ -1 & -3 & -2 & 4 \end{vmatrix}\leqslant 30\cdot(780)^{1/2}\doteq 837.$$

(参阅第七章选做题第12题的(ii)。)

(提示：先证(i)。当A是奇异阵时，不等式显然成立。当A是非异阵时，AA'是正定阵，应用第8题的结论即得(i)。由(i)即得(ii)。)

10. 设$A=(a_{ij})_{m\times n}$是实矩阵，

(i) 如果$A=(B, C)$是A的任意1×2分块阵，求证：

$$|A'A|\leqslant|B'B|\cdot|C'C|;$$

(ii) 求证：$|A'A|\leqslant\prod_{j=1}^{n}\sum_{i=1}^{m}a_{ij}$(Hadamard 不等式的推广)。

11. 设A是正定阵，求证：二次函数$f(x)=x'Ax-2b'x+c$的极小值是$c-b'A^{-1}b$。

12. 设A是负定阵，求$f(x)=x'Ax-2b'x+c$的极大点与极大值。

§ 5 Hermite 型(概述)

由于应用上的需要，前四节讨论的都是实二次型。本节将对

复二次型$\overline{x}'Ax$作概要的叙述，这里A是n阶复方阵，而x是n维复的列向量①，称A为$\overline{x}'Ax$的系数矩阵。由于复方阵A必有唯一分解式：

$$A=H_1+\sqrt{-1}H_2, \tag{1}$$

而$H_1=\dfrac{A+\overline{A}'}{2}$，$H_2=\dfrac{A-\overline{A}'}{2\sqrt{-1}}$都是Hermite阵，故由(1)式可得：

$$\overline{x}'Ax=\overline{x}'H_1x+\sqrt{-1}\overline{x}'H_2x.$$

所以下面主要讨论$\overline{x}'Hx$形状的二次型，其中H是Hermite阵($\overline{H}'=H$)。称$\overline{x}'Hx$为**Hermite型**。当H是实方阵，x是实向量时，Hermite型就是实二次型，所以它也是实二次型的推广。

易知本章前四节关于实二次型与实对称阵的一切结论与证明方法几乎可以相应地搬到Hermite型与Hermite阵上来，所以我们只叙述相应的主要结论，而不再证明它们。

设A是Hermite阵，由于对任何n维复的列向量x，$\overline{x}'Ax$总是实数，故引进

定义 称Hermite型$\overline{x}'Ax$为正定的(或**正定Hermite型**)。如果对任何n维复的列向量$x\neq0$，恒有$\overline{x}'Ax>0$。相应地，称A为**正定Hermite阵**。

定理5.1 Hermite阵是正定的充要条件是，A的特征值全大于零。

定理5.2 设A是正定Hermite阵，则必

(i) $A=\overline{B}'B$，而B是非异复方阵；

(ii) $A=\overline{R}'R$，而R是复的非异上三角阵；

(iii) A的所有主子式全大于零。

定理5.3 如果n阶Hermite阵满足下列条件中的任意一个：

(i) $A=\overline{B}'B$，而B是非异阵；

(ii) A的n个顺序主子式全大于零。则A必是正定Hermite

① 不讨论形如$x'Ax$的二次型，其中A是复方阵，x是复向量。

阵。

定理 5.4 正定 Hermite 阵 A 的实部(矩阵)必是(实对称)正定阵。

【证明】 设 $A=B+\sqrt{-1}C$ 是 A 的分解式，其中 B 与 C 分别是 A 的(实部)矩阵与(虚部)矩阵。由假设 A 是 Hermite 阵，故 B 是实对称阵(C 是反对称阵)。又由假设 A 是正定 Hermite 阵，故对任何 n 维复的列向量 $z\neq 0$，恒有

$$\overline{z}'Az=\overline{z}'(B+\sqrt{-1}C)z>0。\tag{2}$$

今特别取 z 为任一非零实向量 x，则(2)式化为

$$x'Ax=x'Bx+\sqrt{-1}\,x'Cx=x'Bx>0,$$

所以 B 是正定阵。 证毕。

定理 5.4 的推广见之于本节习题第 3 题。

习 题

1. 设 A 是 Hermite 阵，求证：对任一 n 维复的列向量 α，$\overline{\alpha}'A\alpha$ 必是实数。

2. 叙述半正定 Hermite 阵的定义及有关的结论。

3. (Taussky) 设 B 是正定 Hermite 阵 A 的实部(矩阵)，求证：$|B|\geqslant|A|$，而等号成立的充要条件是，A 为(实)正定阵。

4. 设 $A=(a_{ij})_{n\times n}$ 是正定 Hermite 阵，k 是任一正整数，求证：$(\overline{a_{ij}^{k}})_{n\times n}$ 也是正定 Hermite 阵。

§6 双线性型(简介)

设 $x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n$ 为两组实变数，$A=(a_{ij})_{m\times}$ 是实矩阵。记 $x=(x_1, x_2, \cdots, x_m)'$，$y=(y_1, y_2, \cdots, y_n)'$，称

$$f(x, y)=f(x_1, \cdots, x_m; y_1, \cdots, y_n)=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}x_iy_j=x'Ay$$

为 $x_1, \cdots, x_m; y_1, \cdots, y_n$ (或 x, y)的**双线性型**，而称 A 为它的系数矩阵。

当 $m=n$，A 是实对称阵时，称 $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j = x'Ay$ 为**对称双线性型**；当 $m=n$，A 是反对称(实)阵时，称 $x'Ay$ 为**反对称双线性型**。

对称双线性型是通常的(对称)二次型的推广。但须注意，即使 $m=n$，双线性型 $x'Ay$ 却不能像通常的实二次型那样看作对称双线性型，因为：$x_iy_j \neq x_jy_i$，所以它要比实二次型复杂一些。

对一般的双线性型：$x'Ay=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_iy_j$，如用不同的实的非异坐标变换：$x=PZ$，$y=QW$，此处 P 与 Q 分别是 m 阶非异阵与 n 阶非异阵；Z 与 W 分别是 m 维与 n 维实的列向量，则双线性型化为：

$$x'Ay = Z'(P'AQ)W.$$

如果 A 的秩为 r，则必可找到非异阵 P' 与 Q，使 $P'AQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$，于是可得

命题 必存在非异坐标变换 $x=PZ$，$y=QW$，使

$$x'Ay = z_1w_1 + z_2w_2 + \cdots + z_rw_r,$$

此处 $z=(z_1, z_2, \cdots, z_m)'$，$W=(w_1, w_2, \cdots, w_n)'$，$r$ 是 A 的秩。

设 n 阶实对称阵 A 的秩与正惯性指数分别是 r 与 p，则由 §2 (7)式显然可得

命题 2 设 $x'Ay$ 是对称双线性型，则必存在非异坐标变换 $x=Pz$，$y=PW$，使

$$x'Ay = z_1w_1 + z_2w_2 + \cdots + z_pw_p - z_{p+1}w_{p+1} - \cdots - z_rw_r,$$

此处 $z=(z_1, z_2, \cdots, z_n)'$，$W=(w_1, w_2, \cdots, w_n)'$。

由第七章定理 2.2 还显然可得

定理 6.1 设 $x'Ay$ 是对称双线性型，则必存在正交坐标变换 $x=Qz$，$y=QW$，使

$$x'Ay = \lambda_1w_1z_1 + \lambda_2w_2z_2 + \cdots + \lambda_rw_rz_r,$$

其中 Q 是正交阵，r 是 A 的秩，$\lambda_1, \lambda_2, \cdots, \lambda_r$ 是 A 的非零特征值。

当 A 是 Hermite 阵，则称 $\overline{x}'Ay$ 为 **Hermite 双线性型**，此处 x

与 y 都是 n 维复的列向量。易知 Hermite 型的某些结论对相应的 Hermite 双线性型也成立。

习　　题

1. 求证：必存在非异坐标变换：$x=Pz$, $y=PW$, 把反对称双线性型 $x'Ay$ 化为：

$$x'Ay=z_1w_2-z_2w_1+z_3w_4-z_4w_3+\cdots+z_{r-1}w_r-z_rw_{r-1},$$

其中 $2r$ 是反对称阵 A 的秩，$W=(w_1, w_2, \cdots, w_n)'$, $z=(z_1, z_2, \cdots, z_n)'$.

2. 求证：对称双线性型可分解为若干个实二次型之和。从而，对称双线性型由它的二次型唯一决定。

选　做　题

1. 以 $A>0$ 与 $A\geqslant 0$ 分别表示 A 是正定阵与半正定阵，$A>B>0$ 与 $A\geqslant B>0$ 分别表示 A、B、$A-B$ 都是正定阵与 A、B 是正定阵、但 $A-B$ 是半正定阵，求证：

(i) 如 $A>0$, 则 $A+A^{-1}\geqslant 2I(>0)$;

(ii) 如 $A>0$, $B>0$, 则 $|A+B|\geqslant 2|A|^{1/2}|B|^{1/2}$.

2. 设 $A=(a_{ij})_{n\times n}$ 是主对角元全大于零的实对称阵，且

$$\left|\sum_{i=1}^{k}a_{ii}\right|^2>(k-1)\sum_{i,j=1}^{k}a_{ij}^2, \quad k=2, \cdots, n.$$

求证：A 必是正定阵。

（注：此题说明，对满足上述条件的 A，可不必计算其 n 个顺序主子式，而直接得出 A 的正定性）

（提示：应用第七章定理 3.4，先证：$A\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}$ 是非异阵，然后作 $[0, 1]$ 上的连续函数：$f(x)=\left|x[a_{11}, a_{22}, \cdots, a_{kk}]+(1-x)Ac\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}\right|$，由此证明：$\left|A\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}\right|>0$, $2\leqslant k\leqslant n$.）

3. 设 A 是正定阵，S 是反对称（实方）阵，求证：

$$|A+S|>0.$$

（提示：本题可有八种做法。其主要的五种如下：(一) 先证 $|A+S|\neq 0$，再作

$[0, 1]$ 上的连续实值函数 $f(x)=|A+xS|$，说明 $f(1)<0$ 将导至矛盾；(二) 应用第三章 §2 习题第 8 题的 (iii)；(三) 证明 $A+S$ 的特征值的实部全大于零；(四) 直接应用两个方阵之和的行列式的展开式，並应用反对称(实方) 阵的行列式是非负数以及正定阵的性质；(五) 把 A 合同于单位阵 I，再应用反对称的标准形直接推出 $|A+S|>0$；或者用 Schur 定理把 S 面相似于对角阵，即可得证)。

4. (Taussky)设 A 是正定阵，S 是非零反对称(实)方阵，求证：

$$|A+S|>|A|.$$

(这是第 3 题的推广)

5. (i)设 A 是 n 阶实方阵，且 $A+A'$ 是正定阵，求证：

$$2^n|A|\geqslant|A+A'|.$$

当 A 不是实对称阵时，上式的不等号成立。

(ii) 求证：决不存在奇异阵 A，使 $A+A'$ 是正定阵。

6. (Bellman)设 $A=(a_{ij})_{n\times n}$ 是正定阵，$x_1, x_2, \cdots, x_n$ 是实变数，求证：

$$\int_0^\infty\int_0^\infty\cdots\int_0^\infty e^{-\sum_{i,j=1}^{n} a_{ij}x_ix_j}\,dx_1dx_2\cdots dx_n=\frac{\pi^{n/2}}{|A|^{1/2}}.$$

(注：学过多重积分与 Euler 积分后即可做此题)

7. (**华罗庚**)设 A 是实方阵，

(i) 如果 $I-AA'$ 是非异阵，试用行列式的（或矩阵的秩的）第二降阶定理证明：$I-A'A$ 也是非异阵。並证：

$$A'(I-AA')^{-1}=(I-A'A)^{-1}A'.$$

(ii) 如果 B 是实方阵，且 $I-AA'$ 与 $I-BB'$ 是同阶非异阵，求证：

$$\begin{aligned}&(I-AB')(I-BB')^{-1}(I-BA')-(I-AA')\\&=(A-B)(I-B'B)^{-1}(A'-B').\end{aligned}$$

(iii) 如果 $I-AA'$ 与 $I-BB'$ 是同阶正定阵，求证：

$$|I-AA'|\cdot|I-BB'|\leqslant|I-AB'|^2.$$

(iv) 如 $I-AA'$ 与 $I-BB'$ 是同阶正定阵，求证：

$$\begin{pmatrix}(I-AA')^{-1} & (I-AB')^{-1}\\(I-BA')^{-1} & (I-BB')^{-1}\end{pmatrix}$$

是正定阵或半正定阵。

(提示：用正交合同变换把原方阵合同于另一方阵，並应用(iii)。)

(注：本题取自"华罗庚，一个关于行列式的不等式，数学学报(1955) 第 4 期，463～470"中的一部分结论。原文对一般复方阵进行讨论)

8. 设 $A=(a_{ij})_{n\times n}$ 是实方阵，如果A的 n 个顺序主子式全大于零，则称A为拟正阵(例如，正定阵A以及$A+S$都是拟正阵，此处S是反对称(实方)阵。如果A的主对角元全大于零，且

$$\left|\operatorname{tr}\left(A\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}\right)\right|^2>(k-1)\operatorname{tr}\left(A\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}A'\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}\right),$$

$$k=2,\ \cdots,\ n.$$

求证：A 必是拟正阵。

(注：这是第 2 题的推广，证明同第 2 题。)

9. 设A是 n 阶半正定阵（或正定阵），$\lambda_1, \lambda_2, \cdots, \lambda_n$ 是A的特征值，如果正交阵 Q，使 $A=Q[\lambda_1, \lambda_2, \cdots, \lambda_n]Q'$，则记 $A^{1/2}=Q[\lambda_1^{1/2}, \lambda_2^{1/2}, \cdots, \lambda_n^{1/2}]Q'$，求证：

(i) 对满足 $A=P[\lambda_1, \lambda_2, \cdots, \lambda_n]P'$的任何正交阵 P，恒有

$$P[\lambda_1^{1/2}, \lambda_2^{1/2}, \cdots, \lambda_n^{1/2}]P'=f(A),$$

其中 $f(\lambda)$是次数不超过 n 的实系数多项式；

(ii) 存在唯一的半正定阵 B，使 $B^2=A$，从而 $B=A^{1/2}$；

(iii) $A^{1/2}=f(A)$，$(A^{1/2})^2=A$；

(iv) 当A是正定阵时，记 $A^{-1/2}=(A^{1/2})^{-1}$，则

$$A^{-1/2}A^{1/2}=A^{1/2}A^{-1/2}=I,\ (A^{-1/2})^2=A^{-1}.$$

(提示：设 $\lambda_{i1}, \lambda_{i2}, \cdots, \lambda_{is}$ 是 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 中所有不同的特征值，作 Lagrange 内插多项式：

$$f(\lambda)=\sum_{k=1}^{s}\frac{(\lambda-\lambda_{i1})\cdots(\lambda-\lambda_{i_{k-1}})(\lambda-\lambda_{i_{k+1}})\cdots(\lambda-\lambda_{i_s})}{(\lambda_{i_k}-\lambda_{i_1})\cdots(\lambda_{i_k}-\lambda_{i_{k-1}})(\lambda_{i_k}-\lambda_{i_{k+1}})\cdots(\lambda_{i_k}-\lambda_{i_s})}\lambda_{i_k}^{1/2}$$

即可证得(i)式。)

10. 设A与B都是实对称阵，求证：

(i) 如果B是正定阵，则 AB 的特征值全是实数；

(ii) 如果B是正定阵，A 是半正定阵，且 $AB=BA$，则 AB 必是半正定阵；

(iii) 如果A与B都是正定阵，则 AB 的特征值全大于零；

(iv) 两个可交换的正定阵的乘积必是正定阵。

(提示：因为 $AB=(B^{1/2})^{-1}(B^{1/2}AB^{1/2})B^{1/2}$，而 $B^{1/2}AB^{1/2}$是实对称阵，由此即得(i)。)

11. 设A与B是同阶非异实对称阵，求证：A 是正定阵的充要条件是，对所有正定阵 B，恒有 $\operatorname{tr}(AB)>0$。

12. 设A是n阶非异实方阵，求证：必存在n阶正交阵P与Q，使得：$PAQ=[a_1, a_2, \cdots, a_n]$，而$a_i>0$, $i=1, 2, \cdots, n$。

13. 求证：秩为r的$m\times n$实矩阵A必有分解式：

$$A=U\begin{pmatrix}D & 0\\ 0 & 0\end{pmatrix}V', \text{ 而 } D=[a_1, a_2, \cdots, a_r],$$

其中U与V都是正交阵，且$a_i>0$, $i=1, 2, \cdots, r$。称上式为**奇异值分解**。

(提示：先将A写成：$A=M\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}N$，作$M$与$N'$的$QR$分解，然后再应用12题。)

14. 求证：任何n阶实方阵必有分解式：

$$A=(AA')^{1/2}\cdot Q=Q\cdot(A'A)^{1/2},$$

其中Q是正交阵，且当A是非异阵时，Q由A唯一决定。称上式为A的**极因子分解**。

15. (König)设A与B都是实正规阵，P是非异实方阵，且$B=P^{-1}AP$，求证：

(i) $(PP')^{1/2}A=A(PP')^{1/2}$；

(ii) 如果P的极因子分解是$P=(PP')^{1/2}Q$，则$B=Q'AQ$。

(提示：先证存在多项式$f(\lambda)$，使$A'=f(A)$, $B'=f(B)$，然后再证$(P'P)A=A(P'P)$，便可证得(i)。)

16. 求证：任何实方阵必可分解为一个半正定阵与两个对称、正交阵的乘积(提示：应用极因子分解式)。

17. 求证：实方阵A是正规阵的充要条件是，A的极因子分解式为$A=(AA')^{1/2}\cdot Q=Q(AA')^{1/2}$。

18. 设A与B都是n阶实对称阵，且B是正定阵，求证：必存在非异阵P，使

$$P'AP=[\mu_1, \mu_2, \cdots, \mu_n],\ P'BP=I_n,$$

其中$\mu_1, \mu_2, \cdots, \mu_n$必是实数，且是$|A-\lambda B|$的根。

(提示：设对称阵$B^{-1/2}AB^{-1/2}$的特征值是$\mu_1, \mu_2, \cdots, \mu_n$，则存在正交阵$Q$，使$Q'B^{-1/2}AB^{-1/2}Q=[\mu_1, \mu_2, \cdots, \mu_n]$，于是$P=B^{-1/2}Q$即为所求之非异阵，且易证$\mu_1, \mu_2, \cdots, \mu_n$是$|A-\lambda B|$的根。)

19. 设$x'Ax$与$x'Bx$分别是实二次型与正定二次型，求证：必可找到非异坐标变换$x=Py$，把这两个二次型同时化为平方和：

$$x'Ax=y'[\mu_1, \mu_2, \cdots, \mu_n]y,\ x'Bx=y'y,$$

其中 $\mu_1, \mu_2, \cdots, \mu_n$ 是 $|A-\lambda B|$ 的根，它们全是实数。

20. 设 A 与 B 分别是实对称阵与正定阵，求证：必存在实数 μ 及实的列向量 $x \neq 0$，使 $Ax=\mu Bx$，称 μ 为广义特征值(它在应用数学上很有用)。(提示：应用 19 题，或者，应用 A 的 Cholesky 分解也可。)

21. (Ky Fan Taussky)设 A 是正定阵，AB 是实对称阵，求证：AB 是正定阵的充要条件是，B 的特征值全大于零。
(提示：对 AB 与 A 应用第 18 题的结论。)

22. A 与 AB 的假设同 21 题，求证：AB 是半正定(负定)的充要条件是，B 的特征值全是非负数(小于零)。

23. 设 A 与 B 分别是正定阵与半正定阵，如果多项式 $|(1-\mu)A-B|$ 的根 μ 全小于 1，求证：B 必是正定阵。

24. 设 A 是 n 阶正定阵，α 是 n 维实的列向量，求二次型 $x'(A-\alpha\alpha')x$ 的正惯性指数与符号差。

25. 设 A 是 n 阶正定阵，B 是秩为 r 的 n 阶实对称阵，且 $|A-\mu B|$ 的根 λ 全小于 1，求二次型 $x'(A-B)x$ 的正惯性指数与符号差。

26. (Ляпунов) 设 A、C 都是实方阵，且 C 是负定阵，若正定阵 B 使 $AB+BA'=C$，求证：A 的特征值的实部全小于零。(提示：先对 B 与 C 用 18 题找出 P，再应用第七章定理 2.2 的推论 2 于方阵 PAP^{-1}。)
(注：本题之逆见 §4 例 2，例 2 与提示中的方法均取自论文"W. Hahn, Eine Bemerkung Zur zweiten Methode von Ljapunov, Math. Nachr., Bd. 14(1955), s.349～354.")

第九章　线 性 空 间

本章以具体的三维几何空间为雏型，引入抽象线性空间概念，从中可以了解代数抽象的意义与背景；初步掌握用公理系统处理代数问题的思维方法；训练严格逻辑推理的能力。这些不仅是本课程的基本内容之一，而且也是学习近代数学必不可少的。

§1　线性空间的定义

一、线性空间概念的形成

任何抽象概念都是在考察了大量具体事物所具有的共同性质的基础上形成的，线性空间概念也是如此。由于解线性方程组的需要，我们曾把二维平面上和三维空间中的向量推广成一般的 n 维行向量和 n 维列向量，而向量中的每个分量都是某数域 K 中的数。对确定的自然数 n，在由全体 n 维行向量组成的集合 K_n 中，我们曾定义了加法与数量乘法运算如下：若 $k\in K$，而

$$\alpha=(a_1, a_2, \cdots, a_n), \beta=(b_1, b_2, \cdots, b_n)\in K_n,$$

则规定

$$\alpha+\beta=(a_1+b_1, a_2+b_2, \cdots, a_n+b_n),$$

$$k\alpha=(ka_1, ka_2, \cdots, ka_n)=\alpha k,$$

而且证明了这两个运算有如下基本性质（即运算规则）：

(0)：$\alpha+\beta\in K_n$；

(1)：$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$；

(2)：$\theta+\alpha=\alpha+\theta=\alpha$，这里 $\theta=(0, 0, \cdots, 0)$为零向量；

(3)：$\alpha+(-\alpha)=(-\alpha)+\alpha=\theta$，这里 $-\alpha$ 为 α 的负向量；

(0)′：$k\alpha\in K_n$；

(1)′: $k(\alpha+\beta)=k\alpha+k\beta$;

(2)′: $(k+l)\alpha=k\alpha+l\alpha$;

(3)′: $k(l\alpha)=(kl)\alpha$;

(4)′: $1\cdot\alpha=\alpha$;

这里 $\alpha, \beta, \gamma\in K_n$, $k, l\in K$。

实际上，除了 K_n 以外，还存在其他集合，在其中可以定义两种运算，使它也具有上述九条运算规则。例如，设 $M_{m,n}(K)$是数域 K 上的 $m\times n$ 矩阵全体，两种运算就取为通常的矩阵加法和数与矩阵的乘法，则不难一一验证上述九条运算规则都正确，其中 θ 为零矩阵，$-\alpha$ 为 α 的负矩阵。易见，$M_{1,n}(K)=K_n$。今后，特别把 $M_{n,n}(K)$简记为 $M_n(K)$，就是 K 上 n 阶方阵全体。它是本课程中重要的讨论对象之一。

我们发现，不仅在几何和代数中普遍存在着具有这种结构的集合，而且在数学分析中也存在。例如，设 $R[a, b]$表示定义在实闭区间$[a, b]$上的连续实值函数全体，对 $f, g\in R[a, b]$, $k\in K$，定义

$$(f+g)(x)=f(x)+g(x),\ (kf)(x)=kf(x)。$$

又如，$D[a, b]$是$[a, b]$上可微(可导)实函数全体，其中运算同上。易见，在 $R[a, b]$和 $D[a, b]$ 中，上述九条运算规则同样都成立，其中 θ 为取值恒为零的零函数。

在上述这些例子中，集合中的元素分别是向量(或几何空间中的点)，矩阵和函数，它们是些不同的研究对象。但是，可以抽掉这些元素的具体属性，转而考虑某个抽象的集合，在其中定义抽象的"加法"与"数量乘法"，并把上述九条运算规则作为公理。这样，就形成了一个抽象的"代数系统"V。容易明白，在 V 中，凡是仅仅根据这九条公理推导出来的结论，对一切满足这些公理的具体的集合，当然都是正确的。因此，研究这个抽象的代数系统 V，在很大程度上，可代替研究许多具体的互不相同的代数系统，这体现了由具体概括到一般，再由一般指导具体的思维方法的优越性。

现在，我们引入线性空间的概念。

定义 设V是一些元素的非空集合，K是某个数域。如果对V中任意两个元素α,β定义了某种加法，记为$\alpha+\beta$，满足

(A_0)：$\alpha+\beta\in V$，且$\alpha+\beta$由α与β唯一确定，即加法运算的封闭性与单值性；

(A_1)：$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$，即加法运算的结合性；

(A_2)：存在元素$\theta\in V$使得对任何$\alpha\in V$都有$\alpha+\theta=\alpha$；

(A_3)：对于任一$\alpha\in V$，必存在$\delta\in V$使得$\alpha+\delta=\theta$；

这里，$\alpha,\beta,\gamma\in V$。另外，对K中任一数k和V中任一元素α，定义了某种数量乘法，记为$k\alpha$，满足

(M_0)：$k\alpha=\alpha k$是V中元素，且$k\alpha$由k和α唯一确定，即数量乘法的封闭性与单值性；

(M_1)：$k(\alpha+\beta)=k\alpha+k\beta$；

(M_2)：$(k+l)\alpha=k\alpha+l\alpha$；

(M_3)：$k(l\alpha)=(kl)\alpha$；

(M_4)：$1\cdot\alpha=\alpha$；

这里$\alpha,\beta,\gamma\in V$，$k,l\in K$，则称V是**K上的线性空间**（或**向量空间**），称K为**基域**，V中元素称为**向量**。特别，V中满足(A_2)的元素θ称为**零向量**，满足(A_3)的元素δ称为α的**负向量**。

在介绍线性空间的例子之前，我们先证明一个重要事实，即

命题1 设V是K上线性空间，则必满足

$(A_2)'$：存在元素$\theta\in V$使得对任何$\alpha\in V$都有

$$\alpha+\theta=\alpha=\theta+\alpha;$$

$(A_3)'$：对于任一$\alpha\in V$，必存在$\delta\in V$使

$$\alpha+\delta=\theta=\delta+\alpha.$$

【证明】 任意取定$\alpha\in V$，由公理(A_2)和(A_3)可设

$$\alpha+\theta=\alpha,\ \alpha+\delta=\theta,$$

这里δ是V中某个向量。据(A_3)，对于这个δ，又必存在$\delta'\in V$使得$\delta+\delta'=\theta$，因此，再利用(A_1)，(A_2)和(A_3)得

$$\begin{aligned}\delta+\alpha&=(\delta+\alpha)+\theta=(\delta+\alpha)+(\delta+\delta')\\&=\delta+(\alpha+(\delta+\delta'))=\delta+((\alpha+\delta)+\delta')\end{aligned}$$

$$=\delta+(\theta+\delta')=(\delta+\theta)+\delta'=\delta+\delta'=\theta.$$

故$(A_3)'$满足。再由

$$\theta+\alpha=(\alpha+\delta)+\alpha=\alpha+(\delta+\alpha)=\alpha+\theta=\alpha,$$

知$(A_2)'$也满足。 证毕。

这个事实说明：如果把线性空间定义中的(A_2)和(A_3)换为$(A_2)'$和$(A_3)'$，而保持其他公理不变，则这样两个定义是等价的。以后，我们就不区分(A_2)与$(A_2)'$，(A_3)与$(A_3)'$了。不过，要注意的是，它们的等价性的证明是借助于结合性条件(A_1)而完成的。

由线性空间的定义可知，K_n是K上的线性空间，$M_{m,n}(K)$也是K上的线性空间。$R[a, b]$和$D[a, b]$都是R上的线性空间。下面再举几个线性空间的例子。

例 1 $K_n[x]$是数域K上的次数不超过n的多项式全体，再添上零多项式所成的集合，则它关于多项式的加法和数与多项式的乘法，成K上的线性空间。但是，次数等于n的实系数多项式全体，即使再添个零多项式，它关于这两个运算也不成线性空间，因为它对加法并不封闭。例如，$(x^n+x)+(-x^n+2x)=3x$，$n\geqslant 2$。数域K上次数任意的多项式(包括零多项式)全体$K[x]$关于上述两种运算也是K上的线性空间。

例 2 复数域C是实数域R上的线性空间，运算就是复数加法和实数与复数的乘法。当然，实数域R也可看成R上的线性空间。更一般地，如果K是数域F的子域，则F必是K上的线性空间，运算就是F中数的加法和K中数与F中数的乘法。例如，R就是有理数域Q上的线性空间，数域

$$Q(\sqrt{2})=\{a+b\sqrt{2}\mid a, b\in Q\},$$

$$Q(\sqrt{2}, \sqrt{3})=\{a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6}\mid a, b, c, d\in Q\},$$

都是Q上的线性空间。当然，$Q(\sqrt{2}, \sqrt{3})$也是$Q(\sqrt{2})$上的线性空间。要注意的是，子域K不能看作扩域F上的线性空间(这里$K\subset F$)，因为F中的数与K中数相乘未必属于K，即封闭性条件不满足。

在本教程中主要讨论实数域和复数域上的线性空间，今后分别简称为实(线性)空间和复(线性)空间。

二、线性空间的基本性质

如果没有特别说明，今后总设V是某数域K上的线性空间。在线性空间的定义中，我们仅对两个向量$\alpha, \beta \in V$定义了加法，运算结果是V中唯一确定的向量$\alpha+\beta$。如果对α，$\beta, \gamma \in V$定义$\alpha+\beta+\gamma=(\alpha+\beta)+\gamma$，则它也是$V$中唯一确定的向量，且据公理$(A_1)$有

$$\alpha+\beta+\gamma=(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma),$$

这就是说，在$\alpha+\beta+\gamma$中可以任意加括号。那末，关于3个以上的向量相加，情况又是怎样呢？我们有下面的

命题2(加法的一般结合律)　设α_1，α_2，…，α_n是V中任意n个向量，$n \geqslant 3$。定义

$$\sum_{i=1}^{n} \alpha_i=\alpha_1+\alpha_2+\cdots+\alpha_n=(\alpha_1+\cdots+\alpha_{n-1})+\alpha_n,$$

则对任何满足$1 \leqslant k \leqslant n-1$的$k$都有

$$\sum_{i=1}^{n} \alpha_i=(\alpha_1+\cdots+\alpha_k)+(\alpha_{k+1}+\cdots+\alpha_n).$$

这就是说，$\alpha_1+\alpha_2+\cdots+\alpha_n$是$V$中唯一确定的向量，它并不依赖在运算过程中所加括号的方法。

【证明】　对n用归纳法证明。当$n=3$时，就是公理(A_1)所述。设本命题对$n-1$成立，要证对n也成立。事实上，对$1 \leqslant k \leqslant n-2$，有

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i &=(\alpha_1+\cdots+\alpha_{n-1})+\alpha_n & \text{(定义)}\\
&=((\alpha_1+\cdots+\alpha_k)+(\alpha_{k+1}+\cdots+\alpha_{n-1}))+\alpha_n & \text{(归纳假设)}\\
&=(\alpha_1+\cdots+\alpha_k)+((\alpha_{k+1}+\cdots+\alpha_{n-1})+\alpha_n) & \text{(公理 } A_1\text{)}\\
&=(\alpha_1+\cdots+\alpha_k)+(\alpha_{k+1}+\cdots+\alpha_{n-1}+\alpha_n), & \text{(归纳假设)}
\end{aligned}$$

而当$k=n-1$时就是定义所述，所以对$1 \leqslant k \leqslant n-1$都成立。

证毕。

进一步可证

命题 3 （加法交换律）。对任意 $\alpha,\beta\in V$ 有

(A_4): $$\alpha+\beta=\beta+\alpha.$$

【证明】 考虑V中元素$(1+1)(\alpha+\beta)$。依次用公理(M_2)、(M_1)、(M_4)和命题 2 可得

$$\begin{aligned}(1+1)(\alpha+\beta)&=1\cdot(\alpha+\beta)+1\cdot(\alpha+\beta)\\&=(1\cdot\alpha+1\cdot\beta)+(1\cdot\alpha+1\cdot\beta)\\&=(\alpha+\beta)+(\alpha+\beta)=\alpha+\beta+\alpha+\beta.\end{aligned}$$

但也可以由公理(M_1)、(M_2)、(M_4)和命题 2 得到

$$\begin{aligned}&(1+1)(\alpha+\beta)=(1+1)\alpha+(1+1)\beta\\&=(1\cdot\alpha+1\cdot\alpha)+(1\cdot\beta+1\cdot\beta)=(\alpha+\alpha)+(\beta+\beta)\\&=\alpha+\alpha+\beta+\beta.\end{aligned}$$

因此，根据运算的单值性有

$$\alpha+\alpha+\beta+\beta=\alpha+\beta+\alpha+\beta.$$

再据命题 1 中的$(A_3)'$，存在$\gamma,\delta\in V$使$\gamma+\alpha=\theta$，$\beta+\delta=\theta$，则应有

$$(\gamma+\alpha)+(\alpha+\beta)+(\beta+\delta)=(\gamma+\alpha)+(\beta+\alpha)+(\beta+\delta),$$

$$\theta+\alpha+\beta+\theta=\theta+\beta+\alpha+\theta.$$

再据命题 1 中的$(A_2)'$即得$\alpha+\beta=\beta+\alpha$。 证毕。

上述三个命题的证明是一个示范：在证明过程中，必须注意逻辑推理的严密性和正确性，每一步都要有根据，切忌“想当然”，否则，将是错误的证明，甚至得出荒谬的结论。

顺便提及，没有必要把加法交换律作为线性空间的公理要求之一。

命题 4 V中的零向量和任一向量的负向量都是唯一的。

【证明】 设θ和θ'都是V中的零向量，则据$(A_2)'$立得

$$\theta=\theta+\theta'=\theta'.$$

设δ和δ'都是某个$\alpha\in V$的负向量，则

$$\begin{aligned}\delta'&=\theta+\delta' &&(A_2)'\\&=(\delta+\alpha)+\delta' &&(A_3)'\end{aligned}$$

$$=\vartheta+(\alpha+\vartheta') \qquad (A_1)$$

$$=\vartheta+\theta \qquad (A_3)'$$

$$=\vartheta. \qquad (A_2)'$$

证毕。

既然，α 的负向量是唯一确定的，我们就可用一个确定的符号 $-\alpha$ 表示它。这是一个完整的记号，现在还不知道它是否就是 $(-1)\alpha$。利用这个记号，在任一线性空间中就有等式

$$(-\alpha)+\alpha=\theta=\alpha+(-\alpha).$$

利用负向量的唯一性，我们可以在 V 中引入减法运算如下：对于 $\alpha,\beta\in V$，规定

$$\alpha-\beta=\alpha+(-\beta),$$

并称 $\alpha-\beta$ 为 α 与 β 之差。

最后，我们证明向量运算的几个常用性质。

命题 5 对 $\alpha,\beta,\gamma\in V$，$k\in K$ 有

(i) 由 $\alpha+\beta=\alpha+\gamma$ 必可推出 $\beta=\gamma$，即加法消去律成立；

(ii) $0\cdot\alpha=\theta$，这里 0 是 K 中的数零；

(iii) $k\theta=\theta$；

(iv) $(-1)\alpha=-\alpha$；

(v) 如果 $k\alpha=\theta$，则或者 $k=0$ 或者 $\alpha=\theta$。

【证明】 (i) 在等式两边各加上 α 的负向量即得证，

(ii) 因为

$$\alpha+0\cdot\alpha=1\cdot\alpha+0\cdot\alpha=(1+0)\cdot\alpha=1\cdot\alpha=\alpha=\alpha+\theta,$$

所以由加法消去律即得 $0\cdot\alpha=\theta$。

(iii) 因为

$$k\alpha+k\theta=k(\alpha+\theta)=k\alpha=k\alpha+\theta,$$

所以 $k\theta=\theta$。

(iv) 因为

$$\theta=0\cdot\alpha=(1+(-1))\alpha=1\cdot\alpha+(-1)\alpha=\alpha+(-1)\alpha,$$

这说明 $(-1)\alpha$ 也是 α 的负向量。据负向量的唯一性知

$$(-1)\alpha=-\alpha.$$

(v) 如果$k\neq 0$，则 $\frac{1}{k}\in K$，且

$$\alpha=1\cdot\alpha=\left(\frac{1}{k}\cdot k\right)\alpha=\frac{1}{k}(k\alpha)=\frac{1}{k}\theta=\theta.$$

证毕。

习　　题

1. 下列各个集合V关于各自定义的加法和数量乘法是否成为数域K上的线性空间？其中Z, Q, R分别表示整数环，有理数域和实数域。

(i) $V=\{(a_1, a_2, \cdots, a_n)\mid a_i\in Z\}$, $K=Q$,
加法和数乘分别是通常的行向量加法和数与向量的乘法；

(ii) $V=\{(a_1, a_2, \cdots, a_n)\mid a_i\in Q\}$, $K=Q$,
向量加法定义为

$$\begin{aligned}&(a_1, a_2, \cdots, a_n)\oplus(b_1, b_2, \cdots, b_n)\\ =&(a_1+2b_1, a_2+2b_2, \cdots, a_n+2b_n),\end{aligned}$$

数量乘法为

$$k(a_1, a_2, \cdots, a_n)=(ka_1, ka_2, \cdots, ka_n);$$

(iii) $V=\{(a_1, a_2, \cdots, a_n)\mid a_i\in Q\}$, $K=Q$,
采用通常的行向量加法，而数量乘法定义为

$$k\odot(a_1, a_2, \cdots, a_n)=(k^2a_1, k^2a_2, \cdots, k^2a_n);$$

(iv) V为平面上向量全体，$K=R$。取通常的向量加法，而数量乘法定义为

$$k\odot\alpha=\theta,\quad \forall k\in R,\ \alpha\in V;$$

(v) V是平面上不平行于某一取定向量的向量全体所成的集合，$K=R$。取通常的向量加法和数量乘法；

(vi) V为如下三个集合中的某一个，$K=R$，取通常的矩阵加法和数与矩阵的乘法：

$V_1=\{n$ 阶(实)对称阵全体$\}$,

$V_2=\{n$ 阶(实)反对称阵全体$\}$,

$V_3=\{n$ 阶实上三角阵全体$\}$.

2. 设

$$H=\{(z_1, z_2)\mid z_1, z_2 \text{ 为复数}\}.$$

在H中定义

$$(z_1, z_2)=(z_3, z_4) \text{ 当且仅当 } z_1=z_3,\ z_2=z_4;$$

$$(z_1, z_2)+(z_3, z_4)=(z_1+z_3, z_2+z_4);$$
$$k(z_1, z_2)=(kz_1, kz_2), \forall k\in R.$$
求证H是R上的线性空间，称为**实四元数空间**.

3．求证：在线性空间中成立下列各等式：

(i) $-(-\alpha)=\alpha$;

(ii) $-(k\alpha)=(-k)\alpha=k(-\alpha)$;

(iii) $k(\alpha-\beta)=k\alpha-k\beta$.

§2 基与维数

一、线性相关与线性无关

定义 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是V中 n 个向量. 如果存在不全为零的 $k_1, k_2, \cdots, k_n\in K$ 使得

$$\sum_{i=1}^{n} k_i\alpha_i=k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n=\theta,$$

则称 $\alpha_1,\alpha_2,\cdots,\alpha_n$(**对 K**)**线性相关**，否则，称为(**对 K**)**线性无关**，即

$$\sum_{i=1}^{n} k_i\alpha_i=\theta \text{ 当且仅当所有 } k_i=0.$$

由定义易见，单个向量 α 线性无关当且仅当 $\alpha\neq\theta$.

例 1 设 E_{ij} 是这样一个 n 阶方阵，它除了(i, j)位置的元素为 1 以外，其余元素都为 0. 对应于 $i=1, 2, \cdots, n$; $j=1,2,\cdots,n$, 这种 E_{ij} 共有 n^2 个. 显然，

$$\sum_{i,j=1}^{n} k_{ij}E_{ij}=(k_{ij})_{n\times n}=0$$

当且仅当所有 $k_{ij}=0$, 所以这 n^2 个 E_{ij}对K线性无关.

例 2 $K_n[x]$中向量 $1, x, x^2, \cdots, x^n$ 对K线性无关.

例 3 $R[0,3]$中向量 e^x, x^2, x 对R线性无关. 这是因为如果等式

$$k_1e^x+k_2x^2+k_3x=0$$

对于任何 x 都成立的话，取 $x=0$ 就得 $k_1=0$. 再由

$$k_2x^2+k_3x=0,$$

立得 $k_2=k_3=0$。

例 4 复数域 C 中向量 1 和 $\sqrt{-1}$ 对 R 线性无关。

例 5 对任何自然数 n, $R[0,2\pi]$ 中 $2n+1$ 个向量

$$1, \sin x,\ \sin 2x, \cdots,\ \sin nx;\ \cos x,\ \cos 2x, \cdots, \cos nx$$

对 R 线性无关。事实上,若有

$$k_0\cdot 1+k_1\sin x+\cdots+k_n\sin nx+l_1\cos x+\cdots+l_n\cos nx=0, \quad (1)$$

因为对任何正整数 t、u, 总有

$$\int_0^{2\pi}\sin tx\cdot\cos ux\cdot dx=0,$$

而当 $t\neq u$ 时又有

$$\int_0^{2\pi}\sin tx\cdot\sin ux\cdot dx=0,$$

$$\int_0^{2\pi}\cos tx\cdot\cos ux\cdot dx=0,$$

所以对任何 $t=1,2,\cdots,n$, 在(1)式两边先乘以 $\cos tx\cdot dx$, 然后从 0 到 2π 积分即得

$$k_0\int_0^{2\pi}\cos tx\cdot dx+l_t\int_0^{2\pi}\cos^2 tx\cdot dx=0.$$

但

$$\int_0^{2\pi}\cos tx\cdot dx=0, \qquad \int_0^{2\pi}\cos^2 tx\cdot dx=\pi,$$

所以所有 $l_t=0, t=1,2,\cdots,n$。而(1)式实际上是

$$k_0+k_1\sin x+\cdots+k_n\sin nx=0.$$

取 $x=0$ 即得 $k_0=0$, 所以

$$k_1\sin x+k_2\sin 2x+\cdots+k_n\sin nx=0. \quad (2)$$

对任一 $u=1,2,\cdots,n$, 在(2)式两边先乘以 $\sin ux\cdot dx$, 再从 0 到 2π 积分,同理可得

$$k_u\cdot\int_0^{2\pi}\sin^2 ux\cdot dx=0.$$

但

$$\int_0^{2\pi}\sin^2 ux\cdot dx=\pi,$$

故所有 $k_u=0, u=1,2,\cdots,n$。于是证得这 $2n+1$ 个向量对 R 线性无关。

由以上各个例子可以看出，对于各种具体的线性空间中的向量，要用恰当的方法来判断它们是否线性无关。

设 V 是数域 K 上某个抽象线性空间，K_n 是行向量空间。由于它们中的向量组的线性相关性有完全相同的基本性质，而且证明方法也一样，所以我们只写出与第三章 §2 中平行的若干结论而略去其证明。

设 $S=\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$ 为 V 中有限向量组。为简便起见，如果 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关(线性相关)，就称 S 线性无关(线性相关)。我们还用记号 $S\subseteq V$ 表示 S 中元素都是 V 中向量。

命题 1 设 S_1 是 V 中有限向量组，S_2 是 S_1 的子组(即 $S_2\subseteq S_1$)。如果 S_2 线性相关，则 S_1 也线性相关。换句话说，如果 S_1 线性无关，则 S_2 也线性无关。

定义 设 $S=\{\alpha_1,\alpha_2,\cdots,\alpha_r\}\subseteq V, \alpha\in V$。如果存在 $k_1,k_2,\cdots,k_r\in K$ 使得

$$\alpha=k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r,$$

则称 α 是 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的(**K 上的**)**线性组合**，或称 α 可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$(**在 K 上**)**线性表出**，那些 k_i 称为**表出系数**。有时也说，α 是 S 的线性组合，或说 α 可由 S 线性表出。

命题 2 设 $S=\{\alpha_1,\alpha_2,\cdots,\alpha_r\}\subseteq V$, $r\geqslant 2$，则 S 线性相关的充分必要条件是，至少存在某个 α_i 是其余向量的线性组合。

命题 3 设 $S_1=\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$ 为 V 中线性无关组，而 $S_2=\{\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1}\}$ 为 V 中线性相关组，则 α_{r+1} 可由 S_1 线性表出，因而 S_2 中任一向量都可由 S_1 线性表出，而且表出系数组都是唯一确定的。

命题 4 设 $S_1=\{\alpha_1,\cdots,\alpha_p\}$, $S_2=\{\beta_1,\cdots,\beta_q\}$ 和 $S_3=\{\delta_1,\cdots,\delta_r\}$ 都是 V 中有限向量组，如果任一 δ_i 可由 S_2 线性表出，任一 β_i 可由 S_1 线性表出，则任一 δ_i 必可由 S_1 线性表出。

为了使书写和定理的证明简洁起见，我们引进一种形式向量

的记号。设V是数域K上线性空间，$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$中任意$n$个向量，则$(\alpha_1,\alpha_2,\cdots,\alpha_n)$就是由$\alpha_1,\alpha_2,\cdots,\alpha_n$构成的**形式向量**。既然形式向量的分量是一些抽象向量，就有必要规定一些运算。很自然地想到，它应与通常的向量和矩阵的运算一致。

我们规定$(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\beta_1,\beta_2,\cdots,\beta_n)$当且仅当$\alpha_i=\beta_i$，$i=1,2,\cdots,n$。关于形式向量的加法规定为

$$(\alpha_1,\cdots,\alpha_n)+(\beta_1,\cdots,\beta_n)=(\alpha_1+\beta_1,\cdots,\alpha_n+\beta_n). \qquad (1)$$

设$A=(a_{ij})_{n\times m}$为K上矩阵，则规定

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)A$$
$$=\left(\sum_{j=1}^{n}a_{j1}\alpha_j,\ \sum_{j=1}^{n}a_{j2}\alpha_j,\cdots,\ \sum_{j=1}^{n}a_{jm}\alpha_j\right), \qquad (2)$$

所得结果仍是形式向量，只不过含有m个分量罢了。如果把A按它的列分块：

$$A=(\delta_1,\delta_2,\cdots,\delta_m),$$

则(2)式右端形式向量的任一分量可写成

$$\sum_{j=1}^{n}a_{jk}\alpha_j=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}a_{1k}\\a_{2k}\\\vdots\\a_{nk}\end{pmatrix}$$
$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)\delta_k,\quad 1\leqslant k\leqslant m,$$

故得

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)A=(\alpha_1,\alpha_2,\cdots,\alpha_n)(\delta_1,\delta_2,\cdots,\delta_m)$$
$$=((\alpha_1,\cdots,\alpha_n)\delta_1,(\alpha_1,\cdots,\alpha_n)\delta_2,\cdots,(\alpha_1,\cdots,\alpha_n)\delta_m). \qquad (3)$$

据定义式(1)、(2)容易证得

(i) $$(\alpha_1,\alpha_2,\cdots,\alpha_n)(A+B)$$
$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)A+(\alpha_1,\alpha_2,\cdots,\alpha_n)B; \qquad (4)$$

(ii) $$((\alpha_1,\alpha_2,\cdots,\alpha_n)A)C=(\alpha_1,\alpha_2,\cdots,\alpha_n)(AC). \qquad (5)$$

这里A与B是K上任意$n\times m$阵，C是K上任意$m\times p$阵。

命题 5 设A与B都是K上$n\times m$阵，$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$中$n$个线性无关向量，则由

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)A=(\alpha_1,\alpha_2,\cdots,\alpha_n)B,$$

必可推得 $A=B$。

【证明】 记 $C=A-B=(\gamma_1, \gamma_2, \cdots, \gamma_m)$。则由假设及(3)式可得

$$
\begin{aligned}
&(\alpha_1, \alpha_2, \cdots, \alpha_n)(\gamma_1, \gamma_2, \cdots, \gamma_m) \\
&=((\alpha_1, \cdots, \alpha_n)\gamma_1, (\alpha_1, \cdots, \alpha_n)\gamma_2, \cdots, (\alpha_1, \cdots, \alpha_n)\gamma_m) \\
&=(\theta, \theta, \cdots, \theta),
\end{aligned}
$$

于是 $$(\alpha_1, \alpha_2, \cdots, \alpha_n)\gamma_j=\theta, \quad j=1, 2, \cdots, m.$$

但 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关，故由

$$
(\alpha_1, \alpha_2, \cdots, \alpha_n)\gamma_j=(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{nj} \end{pmatrix}
$$

$$
=c_{1j}\alpha_1+c_{2j}\alpha_2+\cdots+c_{nj}\alpha_n=\theta,
$$

可得 $\gamma_j=(c_{1j}, c_{2j}, \cdots, c_{nj})'=(0, 0, \cdots, 0)'$, $j=1, 2, \cdots, m$。这就证明了 $C=0$, $A=B$。 证毕。

定理 2.1 设 $S_1=\{\alpha_1, \alpha_2, \cdots, \alpha_p\}$ 和 $S_2=\{\beta_1, \beta_2, \cdots, \beta_q\}$ 是 V 中两个有限向量组。如果 S_2 中任一向量都是 S_1 的线性组合且 $q>p$，则 S_2 必线性相关。

【证明】 由假设可得

$$
(\beta_1, \beta_2, \cdots, \beta_q)=(\alpha_1, \alpha_2, \cdots, \alpha_p)\begin{pmatrix} k_{11} & k_{21} & \cdots & k_{q1} \\ k_{12} & k_{22} & \cdots & k_{q2} \\ \cdots & \cdots & \cdots & \cdots \\ k_{1p} & k_{2p} & \cdots & k_{qp} \end{pmatrix}
$$

$$
=(\alpha_1, \alpha_2, \cdots, \alpha_p)(\delta_1, \delta_2, \cdots, \delta_q), \tag{6}
$$

这里 $\delta_j=(k_{j1}, k_{j2}, \cdots, k_{jp})'$, $j=1, 2, \cdots, q$，是 K 上 q 个 p 维列向量。因为 $q>p$，据第三章定理 1.2 知 $\delta_1, \delta_2, \cdots, \delta_q$，线性相关，即存在不全为零的 $l_1, l_2, \cdots, l_q$ 使得

$$
(\delta_1, \delta_2, \cdots, \delta_q)\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_q \end{pmatrix}=0,
$$

于是由(5)、(6)两式即得

$$(\beta_1, \beta_2, \cdots, \beta_q)\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_q \end{pmatrix} = \theta.$$

这就证明了 S_2 是线性相关的。　　　　　　　　　　　　证毕。

推论 1　如果 S_2 线性无关，且其中每一个向量都是 S_1 的线性组合，则 $q \leqslant p$。

推论 2　如果 S_1 和 S_2 都线性无关，且 S_1 中任一向量都可由 S_2 线性表出，S_2 中任一向量都可由 S_1 线性表出，则 $p=q$。

二、基与维维

定义　如果在线性空间 V 中存在某 n 个线性无关的向量，但任意 $n+1$ 个向量都是线性相关的，则称 V 为 **n 维线性空间**，称 n 为 V 的**维数**，记为 $d(V)=n$ 或 $\dim(V)=n$。有时，并不关心维数的确切数值，就把 n 维线性空间统称为**有限维线性空间**。如果在 V 中存在任意多个线性无关向量，则称 V 为**无限维线性空间**。

由定义显然可知，如果线性空间 V 中任意 $n+1$ 个向量都线性相关，则 V 必是有限维线性空间，且 $d(V) \leqslant n$。

如何确定有限维线性空间的维数呢？一个常用的判别方法是下面的

定理 2.2　线性空间 V 是 n 维的充分必要条件是

(i) V 中存在 n 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关；

(ii) V 中任一向量都是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合。

【证明】　必要性。由 $d(V)=n$ 的定义知(i)满足。任取 $\beta \in V$，由维数定义知 $\beta, \alpha_1, \alpha_2, \cdots, \alpha_n$ 线性相关，再由命题 3，β 必是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合，故(ii)满足。

充分性。由条件(i)知 V 中至少有 n 个向量 $\alpha_1, \cdots, \alpha_n$ 线性无关。任取 V 中 $n+1$ 个向量 $\beta_1, \cdots, \beta_{n+1}$，由条件(ii)，每一个 β_i 都是 $\alpha_1, \cdots, \alpha_n$ 的线性组合，故由定理 2.1 知 $\beta_1, \cdots, \beta_{n+1}$ 线性相关。

这就证明了 $d(V)=n$。 证毕。

定义 满足定理 2.2 中条件(i)和(ii)的向量组 $S=\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 称为 n 维线性空间 V 的(一个)**基**，而每个 α_i 称为**基向量**。

由定理 2.2 不难看出：n 维线性空间 V 中任何 n 个线性无关的向量都是 V 的基，且 V 中任一基中所含基向量的个数都等于 V 的维数。

下面我们给出关于基与维数的若干例子。

例 6 在例 1 中已证

$$S=\{E_{ij} \mid i, j=1, 2, \cdots, n\}$$

线性无关。显然，任一 $A=(a_{ij})\in M_n(K)$ 都可表为

$$A=\sum_{i,j=1}^{n} a_{ij}E_{ij},$$

所以 S 是 $M_n(K)$ 的基。$d(M_n(K))=n^2$。

例 7 $d(K_n[x])=n+1$。

例 8 易见，复数域 C 是实数域 R 上的 2 维线性空间，R 是 R 上的 1 维线性空间。

例 9 因为对任何自然数 n，多项式 $1, x, x^2, \cdots, x^n$ 对任意数域 K 都是线性无关的，所以 $K[x]$ 是 K 上的无限维线性空间。

例 10 因为对任何自然数 n，$R[0, 2\pi]$ 中向量组

$$1;\ \sin x, \cdots, \sin nx;\ \cos x, \cdots, \cos nx$$

对 R 线性无关，所以 $R[0, 2\pi]$ 是 R 上无限维线性空间。

本章主要讨论有限维线性空间的性质，有时也会涉及一下无限维线性空间，但它的进一步讨论，已越出本书的范围。

习 题

1. 分别取 K 为复数域 C 和实数域 R，问 $M_2(K)$ 中下列向量是否对 K 线性相关：

(i) $A_1=\begin{pmatrix}1&2\\3&4\end{pmatrix}$, $A_2=\begin{pmatrix}-1&1\\1&1\end{pmatrix}$, $A_3=\begin{pmatrix}0&2\\-1&-1\end{pmatrix}$;

(ii) $A_1=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, $A_2=\begin{pmatrix}\sqrt{-1}&0\\0&\sqrt{-1}\end{pmatrix}$,

$$A_3=\begin{pmatrix}0&\sqrt{-1}\\\sqrt{-1}&0\end{pmatrix},\ A_4=\begin{pmatrix}0&-1\\-1&0\end{pmatrix}.$$

2. $R_3[x]$中下列各向量组是否对R线性相关：

(i) $f_1=2x^2-3x+2$, $f_2=x^2-1$, $f_3=x$, $f_4=4$;

(ii) $f_1=1$, $f_2=x-2$, $f_3=x^2-4x+4$, $f_4=x^3-6x^2+12x-8$.

3. 设$R(-\infty,\infty)$是$(-\infty,\infty)$上的连续实函数全体所成的R上无限维线性空间，问其中下列各向量组是否对R线性无关：

(i) $\alpha_1=x^2$, $\alpha_2=\sin x$, $\alpha_3=xe^x$;

(ii) $\alpha_1=3$, $\alpha_2=4\cdot\sin^2 x$, $\alpha_3=\cos 2x$;

(iii) $\alpha_0=1$, $\alpha_1=\cos x$, $\alpha_2=\cos^2 x$, $\cdots$, $\alpha_n=\cos^n x$;

(iv) $\alpha_1=\sin x$, $\alpha_2=\sin 3x$, $\alpha_3=\cos 5x$, $\alpha_4=\cos 7x$.

4. 设$a_1, a_2, \cdots, a_n$是两两互异的实数，试证$R_{n-1}[x]$中向量组

$$f_1=1+a_1x+(a_1x)^2+\cdots+(a_1x)^{n-1},$$
$$f_2=1+a_2x+(a_2x)^2+\cdots+(a_2x)^{n-1},$$
$$\cdots\cdots\cdots\cdots$$
$$f_s=1+a_sx+(a_sx)^2+\cdots+(a_sx)^{n-1},$$

$s\leqslant n$, 对R线性无关。

5. 设$f_1(x)$, $f_2(x)$, $f_3(x)$是$R[x]$中三个互素的多项式，但它们中的任何两个都有非常数的多项式公因子，求证：$f_1(x)$, $f_2(x)$, $f_3(x)$必线性无关。

6. 求实四元数空间H(见§1习题2)的基与维数。

7. 求出下列实线性空间的维数和一个基：

$V_1=\{n$阶(实)对称阵全体$\}$.

$V_2=\{n$阶(实)反对称阵全体$\}$.

$V_3=\{n$阶实上三角阵全体$\}$.

8. 设$C_n=\{(z_1, z_2, \cdots, z_n)\mid z_i\in C\}$, 规定

$$(z_1, z_2, \cdots, z_n)+(u_1, u_2, \cdots, u_n)$$
$$=(z_1+u_1, z_2+u_2, \cdots, z_n+u_n),$$
$$k\cdot(z_1, z_2, \cdots, z_n)=(kz_1, kz_2, \cdots, kz_n),\ k\in R.$$

(i) 求证C_n是R上的线性空间；

(ii) 求出其维数和一个基。

§3　坐标、基变换与坐标变换

设V是数域K上的n维线性空间。任意取定一个基$\alpha_1, \alpha_2, \cdots, \alpha_n$，则对任一$\beta \in V$，必存在$k_1, k_2, \cdots, k_n \in K$使$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$。由于$\alpha_1, \alpha_2, \cdots, \alpha_n$是线性无关的，故这$n$个数$k_1, k_2, \cdots, k_n$是由$\beta$唯一确定的。如同在解析几何中一样，在线性空间中我们也将引入坐标概念。

定义　设$\alpha_1, \alpha_2, \cdots, \alpha_n$是$V$中取定的一个基，则对任一$\beta \in V$，由线性表出式

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$

$$= (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}$$

所唯一确定的有序数组$k_1, k_2, \cdots, k_n$称为β在基$\alpha_1, \alpha_2, \cdots, \alpha_n$下的**坐标**，而称$n$维列向量$(k_1, k_2, \cdots, k_n)'$为$\beta$在基$\alpha_1, \alpha_2, \cdots, \alpha_n$下的**坐标向量**。

易见，同一向量在不同基下的坐标，一般来说是不相同的，这样就需要讨论同一线性空间V中不同的基之间的关系以及同一向量在不同基下的坐标之间的关系。

首先，我们利用向量的坐标给出有关基的两个结果。

定理 3.1　设$\alpha_1, \alpha_2, \cdots, \alpha_n$是线性空间$V$的基，$V$中向量$\beta_i$在此基下的坐标为$c_{1i}, c_{2i}, \cdots, c_{ni}$，$i = 1, 2, \cdots, r$，$r \leqslant n$，则$\beta_1, \beta_2, \cdots, \beta_r$线性无关的充分必要条件是，由$\beta_i$的坐标所成的下列$n \times r$阵

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1r} \\ c_{21} & c_{22} & \cdots & c_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nr} \end{pmatrix}$$

是列满秩阵。

【证明】 由假设得

$$(\beta_1, \beta_2, \cdots, \beta_r) = (\alpha_1, \alpha_2, \cdots, \alpha_n)C,$$

故对任何 r 维列向量 δ 必有

$$(\beta_1, \beta_2, \cdots, \beta_r)\delta = (\alpha_1, \alpha_2, \cdots, \alpha_n)(C\delta). \tag{1}$$

必要性．设 $\beta_1, \beta_2, \cdots, \beta_r$ 线性无关．如果 C 的秩小于 r，则必存在非零 r 维列向量 δ 使 $C\delta = \theta$，于是由(1)式即知 $\beta_1, \beta_2, \cdots, \beta_r$ 线性相关，此为矛盾，故 C 的秩为 r．

充分性．设 C 是列满秩阵．如果有关系式

$$(\beta_1, \beta_2, \cdots, \beta_r)\delta = \theta,$$

则由(1)式可得

$$(\alpha_1, \alpha_2, \cdots, \alpha_n)(C\delta) = \theta,$$

但 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关，必有 $C\delta = \theta$．再由 C 的 r 个列向量是线性无关的，立得 $\delta = \theta$，这就是说 $\beta_1, \beta_2, \cdots, \beta_r$ 线性无关．证毕．

定义 定理 3.1 中所示的 $n \times r$ 阵 C 称为基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 到向量组 $\beta_1, \beta_2, \cdots, \beta_r$ 的**过渡矩阵**．

要注意的是，$\alpha_1, \cdots, \alpha_n$ 到 $\beta_1, \cdots, \beta_r$ 的过渡矩阵指的是 $\beta_1, \cdots, \beta_r$ 经 $\alpha_1, \cdots, \alpha_n$ 线性表出的系数矩阵的转置矩阵．当 $\alpha_1, \cdots, \alpha_n$ 是基时，$\beta_1, \cdots, \beta_n$ 也是基的充分必要条件是，$\alpha_1, \cdots, \alpha_n$ 到 $\beta_1, \cdots, \beta_n$ 的过渡矩阵为 n 阶非异阵，所以定理 3.1 实际上告诉我们，任意给定一个 n 阶非异阵 C，必可由基 $\alpha_1, \cdots, \alpha_n$ 构造出另一个基 $\beta_1, \cdots, \beta_n$，所以任一线性空间中有无限多个基．

为了区别于行向量空间 K_n，我们以后总把数域 K 上的 n 维列向量空间记为 $K^{(n)}$．

例 1 任取 K 上 n 阶非异阵 $A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$，则由 $A = I_n A$ 知

$$(\alpha_1, \alpha_2, \cdots, \alpha_n) = (e_1, e_2, \cdots, e_n)A,$$

因为 $e_1, e_2, \cdots, e_n$ 是 $K^{(n)}$ 的基，所以非异阵 A 的 n 个列向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 也是 $K^{(n)}$ 的基．

例 2 已知 $1, x, x^2, \cdots, x^n$ 是 $R_n[x]$ 的基，易见它到向量组 $1, (x-a), (x-a)^2, \cdots, (x-a)^n$ 的过渡矩阵为单位上三角阵(这

里 a 为某个取定的实数)。显然它是非异阵，所以后者也是 $R_n[x]$ 的基。

用定理 3.1 可以证明如下十分有用的结论。

定理 3.2 对于 n 维线性空间 V 中任意 r 个线性无关的向量 $\beta_1, \beta_2, \cdots, \beta_r$，$1 \leqslant r \leqslant n-1$，在 V 中必存在另外 $n-r$ 个线性无关的向量 $\beta_{r+1}, \cdots, \beta_n$ 使得 $\beta_1, \beta_2, \cdots, \beta_n$ 构成 V 的一个基。

【证明】 任取 V 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$。设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 到 $\beta_1, \beta_2, \cdots, \beta_r$ 的过渡矩阵为

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots c_{1r} \\ c_{21} & c_{22} & \cdots c_{2r} \\ \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots c_{nr} \end{pmatrix},$$

则由定理 3.1 的必要性知 C 的秩为 r。于是由第三章定理 4.1 知存在 $n-r$ 个线性无关 n 维列向量

$$(c_{1j}, c_{2j}, \cdots, c_{nj})', \ j=r+1, \cdots, n,$$

使

$$\begin{pmatrix} c_{11} & \cdots & c_{1r} & c_{1,r+1} & \cdots & c_{1n} \\ c_{21} & \cdots & c_{2r} & c_{2,r+1} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{n1} & \cdots & c_{nr} & c_{n,r+1} & \cdots & c_{nn} \end{pmatrix}$$

为非异阵。令

$$\beta_j = c_{1j}\alpha_1 + c_{2j}\alpha_2 + \cdots + c_{nj}\alpha_n, \ j=r+1, \cdots, n,$$

则再由定理 3.1 的充分性知 $\beta_1, \cdots, \beta_n$ 线性无关，因而是 V 的基。

证毕。

例 3 $M_2(K)$ 中向量

$$A_1=\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A_2=\begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix}$$

在基

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

下的过渡矩阵是

$$C=\begin{pmatrix}1 & 0\\ 2 & 2\\ 3 & -1\\ 4 & 3\end{pmatrix},$$

因$\boldsymbol{C}$的秩为 2，故 A_1，A_2 线性无关．易见 4 阶阵

$$\begin{pmatrix}1 & 0 & 0 & 0\\ 2 & 2 & 0 & 0\\ 3 & -1 & 3 & 0\\ 4 & 3 & 1 & 4\end{pmatrix}$$

是非异阵，故可作矩阵

$$A_3=\begin{pmatrix}0 & 0\\ 3 & 1\end{pmatrix},\quad A_4=\begin{pmatrix}0 & 0\\ 0 & 4\end{pmatrix},$$

使 A_1，A_2，A_3，A_4 构成 $M_2(K)$的基．

现在，要推导出因基变换引起的坐标变换公式．

前面已经提到，V 中同一个向量在不同基下的坐标可能不同．例如 $R_n[x]$中多项式

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$$

在基 $1, x, x^2, \cdots, x^n$ 下的坐标为

$$a_0,\ a_1,\ \cdots,\ a_n,$$

但由 Taylor 展开式

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots$$

$$+\frac{f^{(n)}(a)}{n!}(x-a)^n,$$

知$f(x)$在基 $1, (x-a), (x-a)^2, \cdots, (x-a)^n$ 下的坐标为

$$f(a),\ f'(a),\ \frac{1}{2!}f''(a),\ \cdots,\ \frac{1}{n!}f^{(n)}(a).$$

定理 3.3 设 δ 在基 α_1，α_2，$\cdots$，α_n 下的坐标向量为$(k_1, k_2, \cdots, k_n)'$，在基 β_1，β_2，$\cdots$，β_n 下的坐标向量为$(l_1, l_2, \cdots, l_n)'$．α_1，α_2，$\cdots$，α_n 到 β_1，β_2，$\cdots$，β_n 的过渡矩阵为 C，则

$$(l_1, l_2, \cdots, l_n)' = C^{-1}(k_1, k_2, \cdots, k_n)'. \tag{2}$$

【证明】 由诸假设条件可得

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)C, \tag{3}$$

$$\delta = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = (\beta_1, \beta_2, \cdots, \beta_n)\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$$

$$= [(\alpha_1, \alpha_2, \cdots, \alpha_n)C]\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n)\left[C\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}\right],$$

于是由§2的命题5易得(2)式。 证毕。

我们把(3)式称为**基变换公式**,(2)式称为**坐标变换公式**,它们都是由过渡矩阵C所确定。

习 题

1. 给定$R_n[x]$中两个线性无关向量$f_1 = x+1$, $f_2 = x^n$, 试再找出$n-1$个向量$f_3, \cdots, f_{n+1}$使$f_1, f_2, \cdots, f_{n+1}$构成$R_n[x]$的基.

2. 在$M_2(R)$中找出A_3, A_4使A_1, A_2, A_3, A_4构成一个基,这里

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}.$$

3. 在$M_2(R)$中

(i) 求$A = \begin{pmatrix} -3 & 5 \\ 4 & 2 \end{pmatrix}$ 在基 $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $A_3 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $A_4 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$下的坐标;

(ii) 求基A_1, A_2, A_3, A_4到基$B_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $B_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $B_4 = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ 的过渡矩阵;

(iii) 求A在基B_1, B_2, B_3, B_4下的坐标.

4. 设

$$\alpha_1=\begin{pmatrix}1\\2\\-1\\0\end{pmatrix},\quad \alpha_2=\begin{pmatrix}1\\-1\\1\\1\end{pmatrix},\quad \alpha_3=\begin{pmatrix}-1\\2\\1\\1\end{pmatrix},\quad \alpha_4=\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix}.$$

(i) 证明 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 线性无关;

(ii) 令 $\beta_1=\alpha_1+\alpha_2, \beta_2=\alpha_2+\alpha_3, \beta_3=\alpha_3+\alpha_4, \beta_4=\alpha_1+\alpha_2+\alpha_3$, 求出 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 到 $\beta_1, \beta_2, \beta_3, \beta_4$ 的过渡矩阵 A 和 $\beta_1, \beta_2, \beta_3, \beta_4$ 到 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的过渡矩阵 B;

(iii) 分别求出 $\delta=(1,0,0,0)'$ 在 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 和 $\beta_1, \beta_2, \beta_3, \beta_4$ 下的坐标;

(iv) 哪些 4 维列向量在这两个基下的坐标相同.

5. 给定 $f(x)=4x^3+2x^2+5 \in R_3[x]$.

(i) 求 $f(x)$ 在基 $1, (x-1), (x-1)^2, (x-1)^3$ 下的坐标;

(ii) 求基 $1, (x-1), (x-1)^2, (x-1)^3$ 到基 $1, (x+1), (x+1)^2, (x+1)^3$ 的过渡矩阵;

(iii) 用 $f(x)$ 在上述两个基之间的坐标变换公式求出 $f(x)$ 在基 $1, (x+1), (x+1)^2, (x+1)^3$ 下的坐标;

(iv) 用 $f(x)$ 关于 $x+1$ 的 Taylor 展开式, 求出 $f(x)$ 在基 $1, (x+1), (x+1)^2, (x+1)^3$ 下的坐标.

§4 子空间、生成向量组与线性包

一、线性子空间

我们知道,如果 Π 是三维几何空间 V_3 中任一过原点的平面,则 Π 中向量关于 V_3 中向量加法和数量乘法,显然也成线性空间. 通常就说 Π 是 V_3 的一个子空间. 仿此, 对抽象线性空间也可引入子空间的概念.

定义 设 V 是数域 K 上线性空间, S 是 V 的非空子集, 如果 S 关于 V 的加法和数量乘法也成 K 上线性空间, 则称 S 是 V 的(**线性**)**子空间**.

因为 S 是线性空间 V 的子集, 所以要判定 S 是子空间不必一一重复验证线性空间的九条公理, 而仅需满足两条封闭性公理就

可以了，这就是

定理 4.1 K 上线性空间 V 的非空子集 S 是子空间的充分必要条件是

(i) $\alpha+\beta\in S,\ \forall\alpha,\ \beta\in S$;

(ii) $k\alpha\in S,\ \forall k\in K,\ \alpha\in S$.

与此等价地，这两个条件可合并成

(iii) $k_1\alpha+k_2\beta\in S,\ \forall k_1,\ k_2\in K,\ \alpha,\ \beta\in S$.

【证明】 必要性是显然的，现证充分性。条件(i)就是公理(A_0)。因为 $S\subseteq V$，所以(A_1)在 S 中当然成立。在条件(ii)中取 $k=-1$，有 $-\alpha=(-1)\alpha\in S$。再由条件(i)得知 V 中的零向量 $\theta=\alpha+(-\alpha)\in S$。由 $S\subseteq V$ 知 θ 也是 S 的零向量，所以在 S 中(A_2)和(A_3)也成立。条件(ii)就是(M_0)。因为 $S\subseteq V$，所以(M_1)，(M_2)，(M_3) 和 (M_4) 在 S 中必定成立。这样就证明了 S 是 K 上的线性空间。

{(i)，(ii)}与(iii)之间的等价性，留给读者证明。　　证毕。

任一线性空间 V 都有两个特殊的子空间。一个是由单个零向量所成的子空间，称为**零空间**。我们规定零空间的维数是零。另一个就是 V 本身。我们称这两个特殊的子空间为**平凡子空间**。除此以外的子空间就称为**非平凡子空间**。

如果线性空间 V 是有限维的，根据维数的定义，V 的任何子空间 W 都是有限维的，且 $\mathrm{d}(W)\leqslant\mathrm{d}(V)$。

我们举一些非平凡子空间的例子。

例 1 $R_n[x]$是 $R[x]$的子空间，基域是 R。

例 2 $D[a,b]$是 $R[a,b]$的子空间，基域是 R。

例 3 R 是 C 的子空间，基域是 Q。

例 4 n 阶实对称阵集合 $S_n(R)$是 $M_n(R)$的子空间，基域为R。

例 5 设 A 是 $m\times n$ 实阵，记

$$N=\{x\mid Ax=b,\ x\in R^{(n)}\},$$

这里 $R^{(n)}$是 n 维实列向量空间，则 N 是 $R^{(n)}$的子空间。称 N 是齐次线性方程组 $Ax=0$ 的**解空间**，也称为矩阵 A 的**核空间**。

事实上，任取 $k_1, k_2 \in R$，$\alpha, \beta \in N$ 有

$$A(k_1\alpha + k_2\beta) = k_1 A\alpha + k_2 A\beta = 0,$$

故由定理 4.1 知 N 是 $R^{(n)}$ 的子空间。

二、线性包与生成向量组

本段将介绍构作和描述线性子空间的一个方法。

设 V 是 K 上任一线性空间。任意取定 $\alpha_1, \alpha_2, \cdots, \alpha_s \in V$。记

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) = \left\{ \sum_{i=1}^{s} k_i \alpha_i \mid \forall k_i \in K \right\}.$$

定理 4.2 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是 V 的子空间。

【证明】 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 显然是 V 的非空子集。任取 $k, l \in K$，$\sum_{i=1}^{s} k_i\alpha_i$，$\sum_{i=1}^{s} l_i\alpha_i \in L(\alpha_1, \alpha_2, \cdots, \alpha_s)$，则有

$$k\left(\sum_{i=1}^{s} k_i\alpha_i\right) + l\left(\sum_{i=1}^{s} l_i\alpha_i\right) = \sum_{i=1}^{s} (kk_i + ll_i)\alpha_i \in L(\alpha_1, \alpha_2, \cdots, \alpha_s).$$

于是由定理 4.1 知 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是 V 的子空间。　　证毕。

定义 称 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 为由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 所**生成(张成)的子空间**，或称为 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的**线性包**，称 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 为 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 的**生成向量组**。每个 α_i 称为**生成元**。

由定义易见，$L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是 V 中所有包含 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的子空间中最小者，它似乎紧紧地把 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 包住，起着"最小上界"的作用，故称为线性包。确切地说，它有如下三个特征性质：

(i) $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是 V 的子空间；

(ii) $L(\alpha_1, \alpha_2, \cdots, \alpha_s) \supset \{\alpha_1, \alpha_2, \cdots, \alpha_s\}$；

(iii) V 中任一包含 $\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ 的子空间 S 必包含 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$。

易见，上述第(iii)个特征性质是定理 4.1 的直接推论。

要注意的是，$L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 的生成向量组未必要求线性无

关，因而它的维数未必就是生成元的个数。例如，若 Π 是三维几何空间 V_3 中某个过原点的平面，$\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 是 Π 上两两不共线的 4 个向量，则 $L(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$就是整个平面 Π。显然，$\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 线性相关，而其中线性无关向量的最大个数 2 恰好就是 Π 的维数。一般地有

定理 4.3 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 的维数等于 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 中线性无关向量的最大个数。

【证明】 设 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的某个极大线性无关组，（即它们线性无关，而任一 α_i 都可表为它们的线性组合，）则由定理 2.1 的推论 2，这个 r 是唯一确定的，与极大无关组的选择无关。据极大无关性的定义可知

$$\alpha_i \in L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}),\ i = 1, 2, \cdots, s.$$

但 $L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r})$是线性空间，故

$$\sum_{i=1}^{s} k_i\alpha_i \in L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}),$$

即

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) \subseteq L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}).$$

因为反向包含是显然的，所以

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s) = L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}).$$

但 $L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r})$的维数为 r，所以定理得证。 证毕。

显然，$L(\alpha_1, \alpha_2, \cdots, \alpha_s)$中线性无关向量个数不会超过 s。如果 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的基，则 $V = L(\alpha_1, \alpha_2, \cdots, \alpha_n)$。

例 6 对任一确定的 n，$R[0, 2\pi]$中的向量

$$1, \sin x, \cdots, \sin nx, \cos x, \cdots, \cos nx$$

的线性包

$$L(1, \sin x, \cdots, \sin nx, \cos x, \cdots, \cos nx)$$

是无限维空间 $R[0, 2\pi]$的 $2n+1$ 维子空间，称为三角多项式空间。

例 7 设 A 是秩为 r 的 $m\times n$ 阵。如果 N 是 $Ax=0$ 的解空间，$\beta_1, \beta_2, \cdots, \beta_{n-r}$ 是 $Ax=0$ 的基础解系，则由基础解系的定义得

$$N = L(\beta_1, \beta_2, \cdots, \beta_{n-r}),$$

故 $Ax=0$ 的解空间的维数为 $n-r$。

例 8 设 A 是秩为 r 的 $m\times n$ 阵，把 A 按列分块

$$A=(\alpha_1, \alpha_2, \cdots, \alpha_n),$$

并设 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的某个极大线性无关组.记

$$W=\{Ax \mid x\in R^{(n)}\},$$

这里 $R^{(n)}$ 是 n 维实列向量空间.则

$$W=L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}).$$

【证明】 易见 W 是 $R^{(m)}$ 的子空间。由 $Ae_j=\alpha_j$ 知 $\alpha_j\in W$，$j=1, 2, \cdots, n$。所以

$$L(\alpha_1, \alpha_2, \cdots, \alpha_n)\subseteq W.$$

反之，设 $x=(x_1, x_2, \cdots, x_n)'$，则由

$$Ax=\sum_{i=1}^{n} x_i\alpha_i$$

知

$$W\subseteq L(\alpha_1, \alpha_2, \cdots, \alpha_n),$$

所以 $W=L(\alpha_1, \alpha_2, \cdots, \alpha_n)$。再由定理 4.3 的证明知

$$W=L(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}).$$ 证毕。

在例 5 中已讲过，例 7 中的 N 称为矩阵 A 的核空间。我们再把例 8 中的 W 称为矩阵 A 的**象空间**。上已证得 W 的维数 $\mathrm{d}(W)=\mathrm{r}(A)=r$。由例 7 和例 8 立刻得到关于给定的秩为 r 的 $m\times n$ 阵 A 的重要结论：A 的核空间的维数与 A 的象空间的维数之和等于 A 的列数即

$$\mathrm{d}(N)+\mathrm{d}(W)=A \text{ 的列数}=\mathrm{d}(R^{(n)}).$$

习　题

1．下列 K 上线性空间 V 的非空子集 S 是否构成子空间：

(i) $K=Q$, $V=Q_n$，即有理系数 n 维行向量全体，$S=Z_n$，即整系数 n 维行向量全体；

(ii) $K=R$, $V=M_2(C)$,

$$K=\left\{\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}\middle| a, b\in C\right\}.$$

这里$\bar{a}$是a的共轭复数；

(iii) $K=R, V=M_n(R)$, S为n阶非异阵全体；

(iv) $K=R, V=M_n(R)$, S为n阶反对称阵全体；

(v) $K=R, V=M_n(R)$, S为n阶上(下)三角阵全体；

(vi) $K=R, V=M_n(R)$, S为n阶纯量阵全体.

2. 求出第1题中凡构成子空间的S的基和维数。

3. 求证，对任一合于$1\leqslant r\leqslant n-1$的$r$, n维线性空间V中必存在r维子空间。

4. 若V_1和V_2是V的子空间, $V_1\subseteq V_2$, 证明

(i) 如果$d(V_1)=d(V_2)$, 则$V_1=V_2$；

(ii) 如果$V_1\subseteq V_2$且$V_1\neq V_2$, 则$d(V_1)<d(V_2)$。

5. 设$k_1\alpha_1+k_2\alpha_2+k_3\alpha_3=\theta$且$k_1k_3\neq 0$, 求证

$$L(\alpha_1, \alpha_2)=L(\alpha_2, \alpha_3).$$

6. 求下列线性包$L(\alpha_1, \alpha_2, \alpha_3)$的基与维数。

(i) $\alpha_1=(2, 1, 3, -1)$, $\alpha_2=(-1, 1, -3, 1)$, $\alpha_3=(1, 0, 1, 1)$；

(ii) $\alpha_1=\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $\alpha_2=\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, $\alpha_3=\begin{pmatrix} 1 & 4 \\ 5 & 1 \end{pmatrix}$.

7. 求证n阶非齐次线性方程组

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n=b, \quad \alpha_i, b\in R^{(n)}$$

有解的充分必要条件是

$$L(\alpha_1, \alpha_2, \cdots, \alpha_n)=L(\alpha_1, \alpha_2, \cdots, \alpha_n, b).$$

§5　子空间的和与直接和

本节所引进的子空间的交、和与直接和的概念，是线性空间理论中必不可少的有效工具之一。借助于直接和，可把高维空间化为低维空间来研究。

一、子空间的和与交，直接和

设V_3是三维(几何)空间，Π是任一过原点的平面，在Π上

任意取定两条过原点的互异直线 L_1 和 L_2，则按向量相加的平行四边形法则可证

$$\Pi=\{\alpha_1+\alpha_2|\alpha_1\in L_1,\ \alpha_2\in L_2\}.$$

又如，Π_1 和 Π_2 是 V_3 中过原点的两个互异平面，则也有

$$V_3=\{\alpha_1+\alpha_2|\alpha_1\in\Pi_1,\ \alpha_2\in\Pi_2\}.$$

这就是说，可以把这个 Π 看成是 L_1 与 L_2 之"和"，而 V_3 是 Π_1 与 Π_2 之"和"。这个"和"的概念可以推广到一般线性空间上去。

定义 设 V_1 和 V_2 是 K 上线性空间 V 的两个子空间，则称

$$V_1+V_2=\{\alpha_1+\alpha_2|\alpha_1\in V_1,\ \alpha_2\in V_2\}$$

为 V_1 与 V_2 之和。

易见，$V_1+V_2=V_2+V_1$。

命题 1 V_1+V_2 是 V 的子空间。

【证明】 显然，V_1+V_2 是 V 的非空子集。任取 $\alpha_1+\alpha_2$，$\beta_1+\beta_2\in V_1+V_2$，$k_1$，$k_2\in K$，则

$$\begin{aligned}&k_1(\alpha_1+\alpha_2)+k_2(\beta_1+\beta_2)\\=&(k_1\alpha_1+k_2\beta_1)+(k_1\alpha_2+k_2\beta_2)\in V_1+V_2,\end{aligned}$$

所以 V_1+V_2 是 V 的子空间。 证毕。

两个子空间 V_1 与 V_2 的和 V_1+V_2 是 V 的同时包含 V_1 与 V_2 的子空间中的最小者，它起着"最小上界"的作用。

若 $V_1=L(\alpha_1,\cdots,\alpha_r)$，$V_2=L(\beta_1,\cdots,\beta_s)$，则易证

$$V_1+V_2=L(\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s).$$

这反映了线性包与子空间的和之间的关系。

设 V_1 与 V_2 是 K 上线性空间 V 的两个子空间，则作为集合，它们有交集

$$V_1\cap V_2=\{\alpha|\alpha\in V_1,\ \alpha\in V_2\}.$$

因为 $\theta\in V_1\cap V_2$，所以 $V_1\cap V_2$ 必为 V 的非空子集，且 $V_1\cap V_2=V_2\cap V_1$。

命题 2 $V_1\cap V_2$ 是 V 的子空间。

【证明】 对 α，$\beta\in V_1\cap V_2$，k，$l\in K$，显然有

$$k\alpha+l\beta\in V_1\quad 和\quad k\alpha+l\beta\in V_2,$$

故 $k\alpha+l\beta\in V_1\cap V_2$。　　　　证毕。

类似地，两个子空间 V_1 与 V_2 之交 $V_1\cap V_2$ 是 V 的同时含在 V_1 与 V_2 之中的子空间中的最大者，它起着"最大下界"的作用。

定义　如果 V_1 与 V_2 是 V 的子空间，且 $V_1\cap V_2=\theta$，即 V_1 与 V_2 的交集仅含零向量，则称子空间 V_1 与 V_2 的和 V_1+V_2 为**直接和**，並记为 $V_1+V_2=V_1\oplus V_2$。有时直接用 $V_1\oplus V_2$ 表示 V_1 与 V_2 的直接和。

定理 5.1　设 V_1 与 V_2 是线性空间 V 的子空间。

(i) 如果 $V_1+V_2=V_1\oplus V_2$，则对任一 $\alpha\in V_1+V_2$，必存在唯一的 $\alpha_1\in V_1$，$\alpha_2\in V_2$ 使得 $\alpha=\alpha_1+\alpha_2$，即 V_1+V_2 中的任一向量的这种分解式是唯一的。

(ii) 如果 V_1+V_2 中的零向量的分解式是唯一的，即 $\theta=\theta+\theta$，则 $V_1+V_2=V_1\oplus V_2$。

【证明】 (i) 设 V_1+V_2 是直接和，则 $V_1\cap V_2=\theta$。如果对 $\alpha\in V_1+V_2$，存在 $\alpha_1,\beta_1\in V_1$，$\alpha_2,\beta_2\in V_2$ 使 $\alpha=\alpha_1+\alpha_2=\beta_1+\beta_2$，则

$$\alpha_1-\beta_1=\beta_2-\alpha_2\in V_1\cap V_2=\theta,$$

所以 $\alpha_1=\beta_1$，$\alpha_2=\beta_2$。

(ii) 任取 $\alpha\in V_1\cap V_2$，则 $\alpha\in V_1$ 且 $\alpha\in V_2$，所以 $-\alpha\in V_2$，于是得到 θ 的两种表示法

$$\theta=\theta+\theta=\alpha+(-\alpha),$$

这里 $\theta,\alpha\in V_1$，$\theta,-\alpha\in V_2$，由条件必有 $\alpha=\theta$，故 V_1+V_2 是直接和。　　　　证毕。

定理 5.2　设 V_1 是 n 维线性空间 V 的 r 维非平凡子空间，则存在 V 的 $n-r$ 维子空间 V_2，使 $V=V_1\oplus V_2$。

【证明】 设 $\alpha_1,\cdots,\alpha_r$ 是 V_1 的基。它们当然是 V 中线性无关向量组。由定理 3.2，必存在 $\alpha_{r+1},\cdots,\alpha_n\in V$ 使 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 是 V 的基。取 $V_2=L(\alpha_{r+1},\cdots,\alpha_n)$。由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是线性无关的，故不难推得 $V_1\cap V_2=\theta$，所以 $V=V_1\oplus V_2$。　　证毕。

若 $V=V_1\oplus V_2$，则称 V_1 为 V_2 在 V 中的**补空间**或**余空间**，也称 V_2 为 V_1 在 V 中的补空间或余空间。显然，此时 V_1 与 V_2 互为

补空间或余空间。

二、维数公式

定理 5.3 设 V_1 与 V_2 是线性空间 V 的两个有限维子空间，则有如下维数公式

$$d(V_1)+d(V_2)=d(V_1+V_2)+d(V_1\cap V_2). \tag{1}$$

【证明】 记 $d(V_1)=s$, $d(V_2)=t$, $d(V_1\cap V_2)=p$.

如果 $p=0$，则 $V_1\cap V_2=\theta$. 任取 V_1 的基 $\beta_1,\cdots,\beta_s$ 和 V_2 的基 $\delta_1,\cdots,\delta_t$. 易见，$\beta_1,\cdots,\beta_s,\delta_1,\cdots,\delta_t$ 是 V_1+V_2 中的线性无关向量组，且

$$V_1+V_2=L(\beta_1,\cdots,\beta_s,\delta_1,\cdots,\delta_t).$$

因此，此时维数公式成立。

现设 $p>0$，且 $\alpha_1,\cdots,\alpha_p$ 是 $V_1\cap V_2$ 的一个基。如果 $p=s$ 或 $p=t$，则 $V_1\cap V_2=V_1$ 或 $V_1\cap V_2=V_2$，必有 $V_1+V_2=V_2$ 或 $V_1+V_2=V_1$. 此时，维数公式也成立。所以仅需考虑 $0<p<s$ 且 $0<p<t$ 的情形。由定理 3.2，$V_1\cap V_2$ 的基 $\alpha_1,\cdots,\alpha_p$ 可分别扩充为 V_1 的基

$$\alpha_1,\cdots,\alpha_p,\beta_1,\cdots,\beta_{s-p}$$

和 V_2 的基

$$\alpha_1,\cdots,\alpha_p,\delta_1,\cdots,\delta_{t-p}.$$

若能证明 $d(V_1+V_2)=s+t-p$，则就得到维数公式(1)。为此，考虑向量组

$$\alpha_1,\cdots,\alpha_p;\beta_1,\cdots,\beta_{s-p};\delta_1,\cdots,\delta_{t-p}. \tag{2}$$

首先证明(2)是 V_1+V_2 的生成向量组，事实上，有

$$V_1=L(\alpha_1,\cdots,\alpha_p,\beta_1,\cdots,\beta_{s-p}),$$
$$V_2=L(\alpha_1,\cdots,\alpha_p,\delta_1,\cdots,\delta_{t-p}),$$

则

$$V_1+V_2=L(\alpha_1,\cdots,\alpha_p,\beta_1,\cdots,\beta_{s-p},\delta_1,\cdots,\delta_{t-p}).$$

其次证明(2)线性无关。设

$$\sum_{i=1}^{p} a_i\alpha_i + \sum_{i=1}^{s-p} b_i\beta_i + \sum_{i=1}^{t-p} c_i\delta_i = \theta,$$

则

$$\sum_{i=1}^{p} a_i\alpha_i + \sum_{i=1}^{s-p} b_i\beta_i = -\sum_{i=1}^{t-p} c_i\delta_i. \tag{3}$$

这说明(3)中右端向量属于$V_1 \cap V_2$，所以有

$$-\sum_{i=1}^{t-p} c_i\delta_i = \sum_{i=1}^{p} d_i\alpha_i.$$

但$\alpha_1, \cdots, \alpha_p, \delta_1, \cdots, \delta_{t-p}$是$V_2$的基，所以必有$c_i = 0$，$i = 1, 2, \cdots, t-p$。再据$\alpha_1, \cdots, \alpha_p, \beta_1, \cdots, \beta_{s-p}$是$V_1$的基，由(3)式立得$a_i = 0$，$i = 1, 2, \cdots, p$；$b_i = 0$，$i = 1, 2, \cdots, s-p$，所以向量组(2)线性无关。 证毕。

根据维数公式立刻得到以下诸结果。

定理5.4 设V_1与V_2是线性空间V的有限维子空间，则$V_1 + V_2$是直接和当且仅当

$$d(V_1 + V_2) = d(V_1) + d(V_2).$$

推论 设V_1与V_2是n维线性空间V的子空间。如果$d(V_1) + d(V_2) > n$，则$V_1 \cap V_2$中必含有非零向量。

【证明】 由条件有

$$d(V_1 + V_2) \leqslant n < d(V_1) + d(V_2),$$

所以据定理5.4知$V_1 + V_2$不是直接和，故$V_1 \cap V_2$中必含有非零向量。 证毕。

例 设$R^{(n)}$是实n维列向量空间。任取n阶幂等阵A，则必有分解式

$$R^{(n)} = W \oplus N,$$

其中W是A的象空间，N是A的核空间。

【证明】 可把任一$\alpha \in R^{(n)}$拆写成

$$\alpha = A\alpha + (\alpha - A\alpha).$$

由象空间的定义知$A\alpha \in W$。再由$A^2 = A$得$A(\alpha - A\alpha) = 0$，所以$\alpha - A\alpha \in N$。于是

$$R^{(n)} = W + N, \quad d(W + N) = n.$$

根据§4的例7和例8，得到

$$d(W+N)=d(W)+d(N).$$

所以，由定理5.4知 $R^{(n)}=W\oplus N$。 证毕。

三、无关子空间，有限个子空间的直接和

前面讨论的是两个子空间的和与直接和。现在，要把它们推广到任意有限多个子空间上去。设 $V_1, V_2, \cdots, V_s$ 是 V 的子空间，因为在线性空间 V 中，向量加法是满足结合律和交换律的，所以可归纳地定义**子空间的和**如下：

$$V_1+V_2+V_3=(V_1+V_2)+V_3,$$

$$V_1+V_2+\cdots+V_s=(V_1+\cdots+V_{s-1})+V_s, \quad s\geqslant 3.$$

易见，有限个子空间的和仍是 V 的子空间，且在和式中可以任意加括号，也可以任意调换求和次序。

三个以上的子空间的直接和的定义就比较复杂一些。我们先注意以下两个例子的不同之处。

如果 L_1，L_2，L_3 分别是三维几何空间 V_3 中过原点的三个坐标轴上的向量集合，则它们都是 V_3 的一维子空间，且满足以下条件：

$$L_1\cap L_2\cap L_3=\theta; \tag{4}$$

$$L_1\cap(L_2+L_3)=L_2\cap(L_1+L_3)=L_3\cap(L_1+L_2)=\theta. \tag{5}$$

如果 L_1，L_2，L_3 分别是二维平面 V_2 中过原点的互异三条直线上的向量集合，它们都是 V_2 的一维子空间，且满足以下条件：

$$L_1\cap L_2\cap L_3=\theta; \tag{6}$$

$$L_1\cap L_2=L_2\cap L_3=L_1\cap L_3=\theta. \tag{7}$$

条件(4)与(6)是一致的。由条件(5)显然可推出条件(7)，但反之却不然，所以条件(5)比条件(7)要求更高。更进一步，由条件(5)可推出条件(4)。所以我们将把满足条件(5)作为子空间直接和定义中的条件之一。

命题3 设 V 是数域 K 上线性空间，则 V 中非零向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$，$s\geqslant 2$，线性无关的充分必要条件是，对任意 $1\leqslant i\leqslant s$ 有

$$L(\alpha_i)\cap(L(\alpha_1)+\cdots+L(\alpha_{i-1})+L(\alpha_{i+1})+\cdots+L(\alpha_s))=\theta, \quad (8)$$

这里，$L(\alpha_i)$是单个向量 α_i 的线性包，即由 α_i 生成的一维子空间。

【证明】 必要性。任取

$$\beta\in L(\alpha_i)\cap(L(\alpha_1)+\cdots+L(\alpha_{i-1})+L(\alpha_{i+1})+\cdots+L(\alpha_s)),$$

则存在 $k_i\in K$, $i=1,2,\cdots,s$ 使

$$\beta=k_i\alpha_i=k_1\alpha_1+\cdots+k_{i-1}\alpha_{i-1}+k_{i+1}\alpha_{i+1}+\cdots+k_s\alpha_s.$$

但 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关，所以所有 $k_i=0$，即 $\beta=\theta$。

充分性。设

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=\theta.$$

如果有某个 $k_i\neq0$，则 α_i 可表为 $\alpha_1,\cdots,\alpha_{i-1},\alpha_{i+1},\cdots,\alpha_s$ 的线性组合，所以

$$\alpha_i\in L(\alpha_i)\cap(L(\alpha_1)+\cdots+L(\alpha_{i-1})+L(\alpha_{i+1})+\cdots+L(\alpha_s)).$$

但 $\alpha_i\neq\theta$，这与条件(8)矛盾，所以 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关。证毕。

这说明，非零向量组的线性无关性可用相应子空间的关系式(8)来刻划。仿此，我们引入无关子空间与直接和的概念。

定义 设 $V_1,V_2,\cdots,V_s$ 是 V 的非平凡子空间，$s\geqslant2$。如果对任意 $1\leqslant i\leqslant s$ 都有

$$V_i\cap(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s)=\theta,$$

则称 $V_1,V_2,\cdots,V_s$ 为**无关子空间**。此时，它们的和 $V_1+V_2+\cdots+V_s$ 就称为**直接和**，並记为

$$V_1+V_2+\cdots+V_s=V_1\oplus V_2\oplus\cdots\oplus V_s.$$

有时，直接用 $V_1\oplus V_2\oplus\cdots\oplus V_s$ 表示 $V_1,V_2,\cdots,V_s$ 的直接和。

定理 5.5 设 $V_1,V_2,\cdots,V_s$ 是 K 上线性空间 V 的有限维非平凡子空间，则以下四个条件是等价的：

(i) $V_1+V_2+\cdots+V_s=V_1\oplus V_2\oplus\cdots\oplus V_s$;

(ii) $(V_1+\cdots+V_{i-1})\cap V_i=\theta$, $i=2,3,\cdots,s$;

(iii) $V_1+V_2+\cdots+V_s$ 中任一向量 α 可唯一地表成

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s,\ \alpha_i\in V_i,\ i=1,2,\cdots,s;$$

(iv) $\mathrm{d}(V_1+V_2+\cdots+V_s)=\mathrm{d}(V_1)+\mathrm{d}(V_2)+\cdots+\mathrm{d}(V_s)$.

【证明】 我们采用循环论证法。

(i)⇒(ii)。任取 $2\leqslant i\leqslant s$，显然有

$$\begin{aligned}&(V_1+\cdots+V_{i-1})\cap V_i\\ \subseteq&(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s)\cap V_i=\theta,\end{aligned}$$

故(ii)成立。

(ii)⇒(iii)。设对 $\alpha\in V_1+V_2+\cdots+V_s$ 有

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s=\beta_1+\beta_2+\cdots+\beta_s,$$

这里 $\alpha_i,\beta_i\in V_i$ $i=1,2,\cdots,s$，则

$$\begin{aligned}\alpha_s-\beta_s&=(\beta_1-\alpha_1)+(\beta_2-\alpha_2)+\cdots+(\beta_{s-1}-\alpha_{s-1})\\&\in V_s\cap(V_1+V_2+\cdots+V_{s-1})=\theta,\end{aligned}$$

所以 $\alpha_s=\beta_s$，$\alpha_1+\cdots+\alpha_{s-1}=\beta_1+\cdots+\beta_{s-1}$。再根据(ii)中其他条件可一一证得 $\alpha_i=\beta_i$，$i=1,2,\cdots,s-1$。故 α 表法唯一。

(iii)⇒(iv)。设 $V_1,V_2,\cdots,V_s$ 的基依次为

$$\begin{aligned}&\alpha_{11},\cdots,\alpha_{1n_1};\alpha_{21},\cdots,\alpha_{2n_2};\\&\cdots\cdots\cdots\cdots;\alpha_{s1},\cdots,\alpha_{sn_s}.\end{aligned}\tag{9}$$

则由子空间和的定义，$V_1+V_2+\cdots+V_s$ 中任一向量 α 必可表为这 $n_1+n_2+\cdots+n_s$ 个向量的线性组合。若能证明向量组(9)线性无关，则就证明了(iv)。设

$$\sum_{i=1}^{s}\sum_{j=1}^{n_i}k_{ij}\alpha_{ij}=\theta,$$

记

$$\beta_i=\sum_{j=1}^{n_i}k_{ij}\alpha_{ij},\quad i=1,2,\cdots,s,$$

显然，$\beta_i\in V_i$，$\sum\limits_{i=1}^{s}\beta_i=\theta$。但由条件(iii)知每一 $\beta_i=\theta$，于是

$$\sum_{j=1}^{n_i}k_{ij}\alpha_{ij}=\theta,\quad i=1,2,\cdots,s.$$

但 $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in_i}$ 是 V_i 的基，故所有 $k_{ij}=0$。这就证明了向量组(9)线性无关，因而是 $V_1+V_2+\cdots+V_s$ 的基。

(iv)⇒(i)。仍设 $V_1,V_2,\cdots,V_s$ 的基构成向量组(9)。由子空间和的定义知(9)是 $V_1+V_2+\cdots+V_s$ 的生成元组，即

$$\begin{aligned}&V_1+V_2+\cdots+V_s\\=&L(\alpha_{11},\cdots,\alpha_{1n_1},\alpha_{21},\cdots,\alpha_{2n_2},\cdots,\alpha_{s1},\cdots,\alpha_{sn_s}).\end{aligned}$$

前已证明 $V_1+V_2+\cdots+V_s$ 的维数就是其生成向量组(9)中极大线性无关向量的个数，现在有

$$d(V_1+V_2+\cdots+V_s)=n_1+n_2+\cdots+n_s,$$

故(9)为线性无关组。

对任一 $1\leqslant i\leqslant s$，任取

$$x_i\in V_i\cap(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s),$$

则

$$\begin{aligned}x_i&=\sum_{j=1}^{n_i}k_{ij}\alpha_{ij}\\&=\sum_{j=1}^{n_1}k_{1j}\alpha_{1j}+\cdots+\sum_{j=1}^{n_{i-1}}k_{i-1,j}\alpha_{i-1,j}\\&\quad+\sum_{j=1}^{n_{i+1}}k_{i+1,j}\alpha_{i+1,j}+\cdots+\sum_{j=1}^{n_s}k_{sj}\alpha_{sj}\ .\end{aligned}$$

但(9)是线性无关组，所有 $k_{ij}=0$，于是 $x_i=\theta$，即

$$V_i\cap(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_s)=\theta,$$

$i=1,2,\cdots,s$。这就证明了(i)成立。 **证毕。**

习　题

1. (i) 若 V 是 4 维空间，V_1 和 V_2 分别是 V 的 3 维和 2 维子空间，确定 $V_1\cap V_2$ 的所有可能出现的维数；

(ii) 对 7 维空间 V 的 4 维子空间 V_1 与 5 维子空间 V_2，讨论上述问题.

2. 设 $V_1,V_2,\cdots,V_s$ 是线性空间 V 的子空间，求证

$$d(V_1+V_2+\cdots+V_s)\leqslant d(V_1)+d(V_2)+\cdots+d(V_s).$$

3. 设 V_1,V_2,V_3 是线性空间 V 的子空间，求证

(i) $V_1\cap(V_1\cap V_2+V_3)=(V_1\cap V_2)+(V_1\cap V_3)$；

(ii) $(V_1+V_2)\cap(V_1+V_3)=V_1+(V_1+V_2)\cap V_3$.

4. 若 $M_n(K)$ 是数域 K 上 n 阶阵全体所成的线性空间，S_1 和 S_2 分别是 K 上 n 阶对称阵和 n 阶反对称阵全体所成的子空间，求证 $M_n(K)=S_1\oplus S_2$.

5. 设 $\alpha=(a_1,a_2,\cdots,a_n)'$ 是 $R^{(n)}$ 中给定的一个实向量，且 $\sum_{i=1}^{n}a_i\neq0$，求证

(i) $S=\{x\mid\alpha'x=0,\ x\in R^{(n)}\}$ 是 $R^{(n)}$ 的子空间；

(ii) $T=\{x=(x_1, x_2, \cdots, x_n) \mid x_1=x_2=\cdots=x_n, x\in R^{(n)}\}$ 是 $R^{(n)}$ 的子空间；

(iii) $R^{(n)}=S\oplus T$。

6. 设 $V=V_1\oplus V_2\oplus\cdots\oplus V_s$，其中

$$V_i=V_{i1}\oplus V_{i2}\oplus\cdots\oplus V_{in_i},\quad i=1, 2, \cdots, s,$$

求证

$$V=V_{11}\oplus V_{12}\oplus\cdots\oplus V_{1n_1}\oplus V_{21}\oplus V_{22}\oplus\cdots\oplus V_{2n_2}$$
$$\oplus\cdots\oplus V_{s1}\oplus V_{s2}\oplus\cdots\oplus V_{sn_s}.$$

§6　映射与变换、线性空间的同构

一、映射与变换

在引入抽象映射的概念之前，我们先回顾一下函数的概念。设 $f(x)$ 是定义在实闭区间 $a\leqslant x\leqslant b$ 上的实单值函数。对 $f(x)$，我们也可以这样看：对于 $[a, b]$ 中的任一实数 x，通过函数 f，可以唯一确定一个实数 y 与之对应，这个 y 就是函数值 $f(x)$。我们把这一事实表示为

$$f: x\to y \quad 或 \quad x\xrightarrow{f} y.$$

把函数概念一般化，我们引进

定义　设 S 和 T 是任意两个集合。如果存在某个对应关系，记为 σ，使得对 S 中任一元素 s，在 T 中**必存在唯一的元素 t 与 s 相对应，则称此 σ 是 S 映入 T（内）的（单值）映射**，记为

$$\sigma: s\to t \quad 或 \quad s\xrightarrow{\sigma} t.$$

在映射的运算过程中，又往往把这件事写成 $t=\sigma(s)$。此时，映射就可记为 $s\to\sigma(s)$。称 t 为 s 在 σ 之下的像，称 s 为 t 在 σ 之下的一个**原像**，它并不一定是唯一的。

必须强调指出：某个对应关系 σ 如果是映射的话，它必须满足以下三个条件：S 中任一元素都有像，像必在 T 中，且像是唯一的，这三者不可缺一。

集合 S 在映射 σ 下的**像集**指的是集合

$$\sigma(S)=\{\sigma(s)\mid s\in S\}\subseteq T.$$

我们进一步讨论映射的类型。以下均设σ是S映入T的映射。

(i) 如果$\sigma(S)=T$，也就是说，T中任一元素t在σ之下至少有一个原像$s:\sigma(s)=t$，则称σ是S映到T的**满射**。易见，任意映射σ必是S映到$\sigma(S)$的满射。

(ii) 如果S中任意两个不同的元素s_1，s_2，在T中的像也不同，也就是说，当$\sigma(s_1)=\sigma(s_2)$时必有$s_1=s_2$，则称σ是S映入T的**单射**。

(iii) 如果σ既是满射，又是单射，我们就说σ是S到T的**双射**或**一一对应**。

现在，我们通过一些例子来说明这些概念。

例 1 对任一$\alpha=(a_1, a_2, \cdots, a_n)'\in R^{(n)}$，可定义其长度

$$\|\alpha\|=\sqrt{a_1^2+a_2^2+\cdots+a_n^2},$$

则$\alpha\xrightarrow{\sigma}\|\alpha\|$是$R^{(n)}$映入$R$的映射。因为不同向量可能有相同的长度，所以$\sigma$不是单射。因为$\|\alpha\|$必是非负实数，所以$\sigma$不是$R^{(n)}$映到$R$的满射。当然，$\sigma$是$R^{(n)}$映到非负实数集合上的满射。

例 2 $M_n(R)$是n阶实方阵集合。对任一$A\in M_n(R)$，可求出其行列式$\det A$，则$A\xrightarrow{\sigma}\det A$是$M_n(R)$映入$R$的映射。因为不同的矩阵的行列式可能相同，所以$\sigma$不是单射。因为对任一实数$a$，必存在$n$阶对角阵$D=[1, \cdots, 1, a]$，其行列式为$a$，所以$\sigma$是$M_n(R)$映到$R$的满射。

例 3 取S为闭区间$[0, 1]$，T为实数集合，则$f(x)=x^2$显然是S映入T的单射，但不是满射。

例 4 设Z和E分别表示全体整数集合和全体偶数集合，则$n\xrightarrow{\sigma}2n$是Z映到E的双射。

确切弄清映射的类型，是十分重要的。不同类型的映射，有着本质不同的性质。读者务需弄清概念，熟练掌握，切勿混淆。

定义 S映入S自身内的映射称为S上的**变换**。如果对所有s

$\in S$ 都有 $\sigma(s)=s$，则称 σ 为**恒等变换**，**或单位变换**，常用 1_s 表示，在无需指明 S 时就记为 1^*.

如何定义两个映射相等？设 σ 和 τ 都是 S 映入 T 的映射．如果对任何 $s\in S$ 都有 $\sigma(s)=\tau(s)$，则称 σ 和 τ **相等**，记为 $\sigma=\tau$．否则，称为**不相等**，记为 $\sigma\neq\tau$．显然，定义在两对不同的集合上的映射是不可能相等的．

两个映射如何相乘？哪些映射可以相乘？设 S,T,U 是三个集合，σ 是 S 映入 T 的映射，τ 是 T 映入 U 的映射，则定义

$$(\tau\sigma)(s)=\tau(\sigma(s)),\qquad \forall s\in S.$$

并称 $\tau\sigma$ 是 σ 与 τ 的**乘积**.(注意：$\tau\sigma$ 是先作用 σ，后作用 τ.) 易见，$\tau\sigma$ 是 S 映入 U 的映射．根据这个定义，$\sigma\tau$ 是没有意义的，除非 $U=S$．此时，$\sigma\tau$ 是 T 上的变换，$\tau\sigma$ 是 S 上的变换．因此，任意集合 S 上的任意两个变换都是可以相乘的．

设 ρ，σ 和 τ 依次是 S 映入 T，T 映入 U 和 U 映入 V 的映射，则容易验证映射的乘积是满足结合律的，即

$$(\tau\sigma)\rho=\tau(\sigma\rho).$$

下面我们着重讨论一下双射．设 σ 是 S 映到 T 的双射．因为 σ 是满射，所以对任一 $t\in T$，必存在 $s\in S$ 使 $\sigma(s)=t$．进一步可证这个 s 是由 t 唯一确定的．事实上，如果另有 $s_1\in S$ 使 $\sigma(s_1)=t$，则由 σ 是单射知 $s=s_1$．因此，对任一 $t\in T$，必有唯一的 $s\in S$ 与之对应，这恰好说明，这个新建立起来的对应关系，记为 τ，是 T 映到 S 的双射．根据 τ 的定义，对任一 $s\in S$ 有

$$(\tau\sigma)(s)=\tau(\sigma(s))=\tau(t)=s,$$

这说明 $\tau\sigma=1_s$．此时，我们称 τ 是 σ 的**逆映射**，记为 $\tau=\sigma^{-1}$．于是 $\sigma^{-1}(t)=s$．注意，只有双射才存在逆映射．故双射又称为**可逆映射**．有时，也用记号

$$\sigma: s\longleftrightarrow t,\quad s\in S,\quad t\in T,$$

表示 S 映到 T 的双射．

若 σ 是 S 映到 T 的双射：$\sigma(s)=t$，则 σ 的逆映射 σ^{-1} 是 T 映到 S 的双射：$\sigma^{-1}(t)=s$，且满足

$$\sigma^{-1}\sigma = 1_S \quad 和 \quad \sigma\sigma^{-1} = 1_T。$$

因而 σ 和 σ^{-1} 互为逆映射。

特别，S 映到 S 自身之上的双射 σ 称为**可逆变换**，σ^{-1} 称为 σ 的**逆变换**，此时有关系式

$$\sigma\cdot\sigma^{-1} = \sigma^{-1}\cdot\sigma = 1_S$$

在不会引起混淆时，简记 $1_S = 1^*$，故上式为

$$\sigma\cdot\sigma^{-1} = \sigma^{-1}\cdot\sigma = 1^*。$$

二、线性空间的同构

设 V 是数域 K 上的抽象的 n 维线性空间，$K^{(n)}$ 是 K 上的 n 维列向量空间。在本章前几节的讨论中，我们发现，凡在 $K^{(n)}$ 中成立的许多结论，竟然在 V 中也成立。这个事实告诉我们：在这两者之间，必定存在着非常密切的联系，它就是本段所要讨论的所谓"同构"。对此，我们先作如下分析。

在 V 中任意取定一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$。则对任一列向量 $\alpha = (k_1, k_2, \cdots, k_n)' \in K^{(n)}$，必可唯一确定向量

$$\sigma(\alpha) = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \in V。$$

所以 $\alpha \to \sigma(\alpha)$ 是 $K^{(n)}$ 映入 V 内的单值映射。这个 α 恰好就是 $\sigma(\alpha)$ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标向量。因为 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的基，所以 $K^{(n)}$ 中不同的列向量 α 与 β 决定 V 中不同的向量 $\sigma(\alpha)$ 与 $\sigma(\beta)$，因此，σ 是单射。对任一 $v = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \in V$，必有 $\alpha = (k_1, k_2, \cdots, k_n)' \in K^{(n)}$ 使 $\sigma(\alpha) = v$，所以 σ 是满射。这就是说，σ 是 $K^{(n)}$ 映到 V 的双射。

进一步可证，这个 σ 还具有如下性质：

$$\sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta);$$
$$\sigma(k\alpha) = k\sigma(\alpha),$$

这里 $\alpha, \beta \in K^{(n)}$，$k \in K$。注意：等式左边的加法与数乘是 $K^{(n)}$ 中的运算，而右边的则是 V 中的加法与数乘。我们用相同的符号表示它们，希望不致于引起混淆。设

$$\alpha = (k_1, k_2, \cdots, k_n)', \beta = (l_1, l_2, \cdots, l_n)'$$

是 $K^{(n)}$ 中两个向量，则

$$\alpha+\beta=(k_1+l_1,\ k_2+l_2,\ \cdots,\ k_n+l_n)',$$
$$k\alpha=(kk_1,\ kk_2,\ \cdots,\ kk_n)'.$$

于是显然有

$$\begin{aligned}\sigma(\alpha+\beta)&=(k_1+l_1)\alpha_1+(k_2+l_2)\alpha_2+\cdots+(k_n+l_n)\alpha_n\\&=\sigma(\alpha)+\sigma(\beta),\\ \sigma(k\alpha)&=(kk_1)\alpha_1+(kk_2)\alpha_2+\cdots+(kk_n)\alpha_n\\&=k\sigma(\alpha).\end{aligned}$$

这说明，$K^{(n)}$ 与 V 之间的这个双射 σ 是保持向量加法与数量乘法的一一对应，所以凡是仅仅由加法与数乘推导出来的所有性质，如果在 $K^{(n)}$ 中成立，则经过这种双射 σ 以后，在 V 中一定成立．因此，要知道 V 中有什么结论，只要看一看，在 $K^{(n)}$ 中有什么结论就可以了．由于这个缘故，我们就说，$K^{(n)}$ 与 V 有相同的代数结构，或者说，它们有共同的构造，其确切的定义如下：

定义　设 V 与 V' 是数域 K 上两个线性空间（不一定是有限维的），σ 是 V 映到 V' 的双射．如果 σ 满足

$$\text{(i)}\ \ \sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),\ \forall\alpha,\ \beta\in V;$$
$$\text{(ii)}\ \ \sigma(k\alpha)=k\sigma(\alpha),\ \forall\alpha\in V,\ k\in K, \tag{1}$$

或者与此等价地，σ 满足

$$\text{(iii)}\ \ \sigma(k\alpha+l\beta)=k\sigma(\alpha)+l\sigma(\beta),\ \forall\alpha,\ \beta\in V,\ k,\ l\in K.$$

则称 σ 是 V 到 V' 上的**同构映射**或 **K—同构**．特别，V 到 V 自身之上的同构映射称为 **K—自同构**．

称 V 与 V' 是同构的，如果至少存在一个 V 到 V' 上的同构映射 σ．此时记为 $V\overset{\sigma}{\cong}V'$．若并不关心具体的 σ，则可简记为 $V\cong V'$．

若 $V\overset{\sigma}{\cong}V'$，则 σ 是 V 到 V' 上的同构映射，且满足(1)式．因为 σ 是双射，其逆映射 σ^{-1} 是 V' 到 V 的双射，进一步有

$$\begin{aligned}\sigma^{-1}(\sigma(\alpha)+\sigma(\beta))&=\sigma^{-1}(\sigma(\alpha+\beta))=\alpha+\beta\\&=\sigma^{-1}(\sigma(\alpha))+\sigma^{-1}(\sigma(\beta));\\ \sigma^{-1}(k\sigma(\alpha))&=\sigma^{-1}(\sigma(k\alpha))=k\alpha=k\sigma^{-1}(\sigma(\alpha)).\end{aligned}$$

这里 $\sigma(\alpha)$和 $\sigma(\beta)$是 V' 中任意两个向量(这是因为 σ 是满射,V'中任一向量必可写为$\sigma(\alpha)$的形状),k 是 K 中任意数。这说明 σ^{-1} 是 V' 到 V 上的同构映射,即$V' \overset{\sigma^{-1}}{\cong} V$。所以,线性空间之间的同构关系具有对称性,即当 $V \cong V'$ 时,必有 $V' \cong V$,反之也然。更进一步,它还具有传递性:若 $V \cong V'$, $V' \cong V''$,则 $V \cong V''$。事实上,设 σ 和 τ 分别是 V 到 V' 和 V' 到 V'' 上的同构映射,则不难验证双射 σ 和 τ 的乘积 $\tau\sigma$ 仍是双射,且仍满足(1)式,所以 $\tau\sigma$ 是 V 到 V''上的同构映射。显然,同构关系具有反身性:$V \cong V$,其同构映射可取为恒等映射。因此,我们可把 K 上线性空间全体按同构关系分类。

根据本段开始的讨论,立即可得

命题 1 数域 K 上的任一 n 维线性空间 V 都与 $K^{(n)}$同构,即 $V \cong K^{(n)}$。

另外,线性空间之间的同构映射还有如下性质。

命题 2 设 σ 是 V 到 V' 上的同构映射,则

(i) $\sigma(\theta)=\theta'$,这里 θ 和 θ' 分别是 V 和 V' 的零向量;

(ii) $\sigma(-\alpha)=-\sigma(\alpha)$, $\forall \alpha \in V$;

(iii) $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 中线性无关向量组的充分必要条件是 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)$是 V' 中线性无关向量组。

【证明】 (i) 任取 $\alpha \in V$,有

$$\sigma(\theta)=\sigma(0\cdot\alpha)=0\cdot\sigma(\alpha)=\theta'.$$

(ii) $\sigma(-\alpha)=\sigma((-1)\alpha)=(-1)\sigma(\alpha)=-\sigma(\alpha)$。

(iii) 设

$$k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_n\sigma(\alpha_n)=\theta',$$

因为 σ 是同构映射,用归纳法可得

$$\sigma\left(\sum_{i=1}^{n} k_i \alpha_i\right)=\sum_{i=1}^{n} k_i\sigma(\alpha_i)=\theta',$$

但 $\sigma(\theta)=\theta'$,σ 是单射,所以 $\sum\limits_{i=1}^{n} k_i\alpha_i=\theta$。但 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关,所有 $k_i=0$,所以 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)$线性无关。

反之，设 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)$线性无关。如果 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性相关，存在不全为零的 $k_1, k_2, \cdots, k_n$ 使 $\sum_{i=1}^{n} k_i\alpha_i = \theta$，则有

$$\sum_{i=1}^{n} k_i\sigma(\alpha_i) = \sigma\left(\sum_{i=1}^{n} k_i \alpha_i\right) = \sigma(\theta) = \theta',$$

这是一个矛盾，所以 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 必线性无关。 **证毕。**

定理 6.1 设 V 和 V' 是数域 K 上两个有限维线性空间，则 $V \cong V'$ 当且仅当 $d(V) = d(V')$。

【证明】 必要性。设 σ 是 V 到 V' 上的同构映射，$d(V) = n$。任取 V 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$。由命题 2，$\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)$是 V' 中线性无关组。任取 $\beta' \in V'$，因为 σ 是满射，存在 $\beta \in V$ 使 $\sigma(\beta) = \beta'$，又

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$

在同构映射 σ 之下有

$$\beta' = \sigma(\beta) = k_1\sigma(\alpha_1) + k_2\sigma(\alpha_2) + \cdots + k_n\sigma(\alpha_n).$$

这就证明了 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)$是 V' 的基。因此 $d(V') = n = d(V)$。

充分性。设 $d(V) = d(V') = n$。由命题 1 有 $V \cong K^{(n)}$，$V' \cong K^{(n)}$。由同构关系的对称性得 $K^{(n)} \cong V'$，再由传递性得 $V \cong V'$。

证毕。

习　题

1. 确定下列从集合 S 映入集合 T 的映射 σ 的类型(单射、满射或双射)：

(i) $S = Q_n[x]$, $T = R$, $\sigma: f(x) \to \int_0^1 f(x)dx$;

(ii) $S = Q_n[x]$, $T = Q$, $\sigma: f(x) \to \int_0^1 f(x)dx$;

(iii) $S = T = \{x | 0 \leqslant x \leqslant 1, x \in R\}$, $\sigma: x \to x^3$;

(iv) $S = K_n$, $T = K$, $\sigma: (a_1, a_2, \cdots, a_n) \to \sum_{i=1}^{n} a_i$;

(v) $S = R_n$, $T = R$, $\sigma: (a_1, a_2, \cdots, a_n) \to \sum_{i=1}^{n} a_i^2$;

(vi) $S = M_n(R)$, $T = C_n$,

$$\sigma: \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \to (a_{11}, a_{12}, \cdots, a_{1n});$$

(vii) $S = T = R_n[x]$, $\sigma: f(x) \to f'(x)$, 就是求 $f(x)$ 的导数；

(viii) $S = \{(x, y) \mid x, y \in R, x^2 + y^2 = 1\}$,

$T = \{(x_1, y_1) \mid x_1, y_1 \in R, x_1^2 + y_1^2 = 4\}$,

$\sigma: (x, y) \to (x_1, y_1) = (2x, 2y)$.

以上出现的 Q, R 和 C 依次表示有理数域、实数域和复数域，而 K 是任一数域。

2. 应用命题 1 的结论证明以下事实．设 V 是数域 K 上 n 维线性空间，$\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的基，$S = \{\beta_1, \beta_2, \cdots, \beta_m\}$ 是 V 中有限向量组。如果

$$\beta_i = a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n, \; i = 1, 2, \cdots, m,$$

令 $A = (a_{ij})_{m\times n}$, 则 S 中极大线性无关向量个数恰好就是 A 的秩。

3. 设 $S = \{aI_n \mid a \in R\}$, 这里 I_n 是 n 阶单位阵．证明 $\sigma: aI_n \to a$ 是 S 到 R 的双射，并写出 σ^{-1}, $\sigma\cdot\sigma^{-1}$ 和 $\sigma^{-1}\cdot\sigma$。

选 做 题

1. $C_n[x]$ 表示次数不超过 n 的复系数多项式集合添上零多项式所成的复线性空间(取通常的多项式加法与数乘)．设

$$\omega_k = e^{\sqrt{-1}\cdot\frac{2k\pi}{n}}, \; k = 1, 2, \cdots, n,$$

(i) 求证 $C_{n-1}[x]$ 中的量

$$f_i(x) = (x - \omega_1)\cdots(x - \omega_{i-1})\,(x - \omega_{i+1})\cdots(x - \omega_n), \; i = 1, 2, \cdots, n$$

对 C 线性无关；

(ii) 求基 $1, x, x^2, \cdots, x^{n-1}$ 到基 $f_1, f_2, \cdots, f_n$ 的过渡矩阵；

(iii) 求向量 $f(x) = \omega_1 + \omega_2 x + \omega_3 x^2 + \cdots + \omega_n x^{n-1}$ 在基 $f_1, f_2, \cdots, f_n$ 下的坐标。

2. 求证：数域 K 上 n 维列向量空间 $K^{(n)}$ 的任何一个非平凡子空间 N 都可作为某个齐次线性方程组 $Ax = 0$ 的解空间。

3. 设 A 是秩为 r 的 n 阶半正定阵，$R^{(n)}$ 是 n 维实列向量空间，求证

$$V_1 = \{\alpha \mid \alpha' A\alpha = 0, \; \alpha \in R^{(n)}\}$$

是 $R^{(n)}$ 的 $n - r$ 维子空间。

4. 设$\delta\neq 0$是满秩实二次型$f(x_1, x_2, \cdots, x_n)=x'Ax$的符号差。求证：$R^{(n)}$中必存在$\frac{1}{2}(n-|\delta|)$维子空间$V_2$，使对任何$\alpha\in V_2$必有$\alpha'A\alpha=0$。

5. (Steinitz 替换定理) 设$\beta_1, \beta_2, \cdots, \beta_t$是线性包$L(\alpha_1, \alpha_2, \cdots, \alpha_s)$中的线性无关向量组，$t\leqslant s$，则在$\alpha_1, \alpha_2, \cdots, \alpha_s$中必存在$\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_{s-t}}$，使

$$L(\alpha_1, \alpha_2, \cdots, \alpha_s)=L(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_{s-t}}, \beta_1, \beta_2, \cdots, \beta_t).$$

6. 设V是数域K上的线性空间，求证

(i) 若V_1与V_2是V的两个非平凡子空间，则必存在$\alpha\in V$使$\alpha\notin V_1$且$\alpha\notin V_2$；

(ii) 若$V_1, V_2, \cdots, V_s$是V的s个非平凡子空间，则必存在$\alpha\in V$使$\alpha\notin V_i$，$i=1, 2, \cdots, s$。

(提示：对s用归纳法。作归纳假设：存在$\alpha\in V$使$\alpha\notin V_i$，$i=1, 2, \cdots, s-1$。分两种情形。若$\alpha\notin V_s$，则结论得证。若$\alpha\in V_s$，则由于V_s是非平凡的，可取$\beta\in V$，$\beta\notin V_s$，再就β属于$V_1, V_2, \cdots, V_{s-1}$的某些$V_i$以及$\beta$不属于$V_1, V_2, \cdots, V_{s-1}$中的某些$V_j$进行讨论。在前者情形下，易证$\alpha+k\beta\notin V_i$，$i=1, 2, \cdots, s$；在后者情形下，可证只能有有限多个$k\in K$使$\alpha+k\beta\in V_j$。但$K$是无限域，故必存在无限多个$l$使$\alpha+l\beta\notin V_i$，$i=1, 2, \cdots, s$。)

第十章　线性映射与线性变换

在上一章中已经提到，线性空间之间的同构，刻划了这两个空间的某种本质上的一致，或者说，它们有相同的代数结构。线性空间之间的同构映射首先要求是双射。然而有大量的线性空间之间的映射存在，它们虽然不是双射，但也保持加法与数量乘法的对应关系。这种映射将被称为线性映射。它们是更广泛的，因而也是最基本的映射。本章将讨论线性映射的基本性质，线性变换的基本运算，并把线性变换与线性空间的不变子空间这两个概念紧密联系起来。

§1　线性映射、线性变换与线性函数

定义　设 V 与 V' 是数域 K 上的两个线性空间，σ 是 V 映入 V' 的映射。如果 σ 满足

$$\begin{aligned}&\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),\ \forall\alpha,\beta\in V,\\&\sigma(k\alpha)=k\sigma(\alpha),\qquad\qquad \forall\alpha\in V,\ k\in K,\end{aligned}\tag{1}$$

则称 σ 是**线性映射**或**线性算子**。特别，V 映入 V 自身的线性映射称为**线性变换**。V 映入基域 K 的线性映射称为**线性函数**。

易见，条件(1)等价于条件

$$\sigma(k\alpha+l\beta)=k\sigma(\alpha)+l\sigma(\beta),\tag{2}$$

对任意 $\alpha,\beta\in V,k,l\in K$ 都成立。

显然，两个线性空间之间的同构映射都是线性映射，自同构映射都是线性变换。下面我们举出一些未必是双射的线性映射的例子。

例 1　设 $A=(\alpha_1,\alpha_2,\cdots,\alpha_n)\in M_{m,n}(R)$，则对任意取定的 j：

$1 \leqslant j \leqslant n$，映射 $A \xrightarrow{\sigma} \alpha_j$ 就是 $M_{m,n}(R)$ 到 $R^{(m)}$ 上的线性映射，即 σ 是满射，但不是单射。

例 2 多项式的微分运算 $f(x) \to f'(x)$ 是 $R[x]$ 上的线性变换。这是因为对 $f(x), g(x) \in R[x], k \in R$ 有

$$(f(x)+g(x))' = f'(x)+g'(x),\ (kf(x))' = kf'(x),$$

但它不是单射。

例 3 $R[a, b]$ 是 $[a, b]$ 上实连续函数全体，则 $f(x) \xrightarrow{\sigma} \int_a^b f(x) dx$ 是 $R[a, b]$ 到 R 上的线性函数。这是因为，σ 是单值函数，且对任何 $f(x), g(x) \in R[a, b], k, l \in R$ 都有

$$\int_a^b (kf(x)+lg(x))\mathrm{d}x = k\int_a^b f(x)\mathrm{d}x + l\int_a^b g(x)\mathrm{d}x.$$

显然，σ 不是单射。

例 4 取定 $B \in M_n(K)$。令

$$\sigma: A \to AB,\ \forall A \in M_n(K).$$

它是 $M_n(K)$ 映入自身的单值映射，且

$$\sigma(A_1+A_2) = \sigma(A_1)+\sigma(A_2),\ \sigma(kA_1) = k\sigma(A_1)$$

对所有 $A_1, A_2 \in M_n(K), k \in K$ 都成立，所以它是线性变换。但它既不是单射，又不是满射，因而不是 $M_n(K)$ 的 K-自同构。

例 5 对 $\alpha = (a_1, a_2, \cdots, a_n) \in R_n$，令 $f(\alpha) = \sum_{i=1}^n a_i$，则 f 是 R_n 上的线性函数。因为对任意 $\alpha, \beta \in R_n, k, l \in R$ 有

$$f(k\alpha+l\beta) = kf(\alpha)+lf(\beta).$$

由线性映射的定义不难证得

命题 1 设 V 和 V' 是数域 K 上的两个线性空间，则 V 映入 V' 的映射 σ 是线性映射的充分必要条件是

$$\sigma\left(\sum_{i=1}^s k_i\alpha_i\right) = \sum_{i=1}^s k_i\sigma(\alpha_i),$$

对于任何自然数 s，任何 $\alpha_1, \alpha_2, \cdots, \alpha_s \in V$ 和 $k_1, k_2, \cdots, k_s \in K$ 都成立。

命题 2 设 σ 是线性空间 V 映入线性空间 V' 的线性映射，

$\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 V 中线性无关组。如果 σ 是单射，则 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，…，$\sigma(\alpha_s)$ 也是 V' 中线性无关组。

注意：如果线性映射 σ 不是单射，则命题 2 未必成立。例如，考虑把任一 $\alpha \in V$ 都映成 V' 中零向量 θ' 的映射

$$O^*: \alpha \to \theta',$$

它当然是线性映射，称为**零映射**。当 $V'=V$ 时，称 O^* 为**零变换**。对零映射 O^*，命题 2 显然不正确。

习 题

1. 判断下列映射 σ 是不是线性映射？

(i) 取定某个 $B \in M_{m,n}(R)$，

$$\sigma: A \to A^2 B, \forall A \in M_m(R);$$

(ii) 对 $A=(a_{ij}) \in M_n(K)$，

$$\sigma: A \to \mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

2. 判定下列变换是不是线性变换？

(i) 将平面上任一向量 α 变为与 α 同一方向上的单位向量；

(ii) 在 K 上 n 维线性空间 V 中取定基 $\alpha_1, \alpha_2, \cdots, \alpha_n$，取定自然数 $m<n$。投影(射影)变换

$$P: \alpha = \sum_{i=1}^{n} k_i \alpha_i \to \alpha' = \sum_{i=1}^{m} k_i \alpha_i;$$

(iii) 取定 $A, B \in M_n(R)$，

$$\sigma: X \to AXB, \forall X \in M_n(R);$$

(iv) 取定 $z_0 \in C$，

$$\sigma: z \to z_0 \cdot z, \forall z \in C;$$

(v) $\sigma: z \to \bar{z}, \forall z \in C$，$\bar{z}$ 表示 z 的共轭复数；

(vi) 对 $\alpha \in R_3$，以 $|\alpha|$ 表示 α 的长度

$$\sigma: \alpha \to \begin{cases} \dfrac{1}{|\alpha|}\alpha, & \text{若 } \alpha \neq \theta, \\ 0, & \text{若 } \alpha = \theta; \end{cases}$$

(vii) $\sigma: f(x) \to \int_a^x f(t)\,dt, \forall f(x) \in R[a, \infty]$；

(viii) 设 $k(x, t)$ 是定义在正方形 $a \leqslant x \leqslant b$，$a \leqslant t \leqslant b$ 上的某个取定的

实连续函数，对 $f(x)\in R[a,b]$，定义

$$\sigma:\ f(x)\to\int_a^b k(x,t)f(t)\mathrm{d}t.$$

称 σ 为 Fredholm 算子；

(ix) 取定 $p(x), q(x)\in R[a,b]$，对 $f(x)\in R[a,b]$，定义

$$\sigma:\ f(x)\to\frac{\mathrm{d}}{\mathrm{d}x}\left(p(x)\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right)+q(x)f(x),$$

称 σ 为 Sturm-Liouville 算子。

3. 证明下列映射都是 $M_n(K)$ 上的线性函数.

(i) $A=(a_{ij})_{n\times n}\to\sum_{i,j=1}^{n}a_{ij}$；

(ii) $A=(a_{ij})_{n\times n}\to\sum_{k=1}^{n}ka_{kk}$。

4. 取定 $A\in M_{m,n}(K)$，

(i) 求证 $x\to Ax$ 是 $K^{(n)}$到$K^{(m)}$的线性映射；

(ii) 再取定 $x_0\in K^{(m)}$，定出

$$\sigma:\ x\to Ax+x_0,\qquad \forall x\in K^{(n)}$$

是线性映射的充分必要条件。

§2 线性映射的矩阵表示

本节中仅讨论数域 K 上的有限维线性空间。

一、线性映射的矩阵表示

设 V 与 V' 分别是 n 维与 m 维线性空间，σ 是 V 映入 V' 的线性映射。任意取定 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，则对任一 $\alpha=\sum_{i=1}^{n}k_i\alpha_i$，由于 $\sigma(\alpha)=\sum_{i=1}^{n}k_i\sigma(\alpha_i)$成立，所以$\sigma(\alpha)$由基的像$\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)$完全确定。在 V' 中任意取定基 $\alpha_1',\alpha_2',\cdots,\alpha_m'$，则对于 $\sigma(\alpha_i)\in V'$ 必有

$$
\begin{aligned}
\sigma(\alpha_1) &= a_{11}\alpha_1' + a_{12}\alpha_2' + \cdots + a_{1m}\alpha_m', \\
\sigma(\alpha_2) &= a_{21}\alpha_1' + a_{22}\alpha_2' + \cdots + a_{2m}\alpha_m', \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\sigma(\alpha_n) &= a_{n1}\alpha_1' + a_{n2}\alpha_2' + \cdots + a_{nm}\alpha_m'.
\end{aligned} \tag{1}
$$

如同在第九章§2中所说的那样，可把(1)式用形式向量写出：

$$(\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)) = (\alpha_1', \alpha_2', \cdots, \alpha_m')A,$$

其中

$$A = \begin{pmatrix} a_{11} & a_{21}\cdots a_{n1} \\ a_{12} & a_{22}\cdots a_{n2} \\ \cdots\cdots & \cdots\cdots \\ a_{1m} & a_{2m}\cdots a_{nm} \end{pmatrix}_{m\times n}.$$

我们把这个A称为线性映射σ在基$\alpha_1,\alpha_2,\cdots,\alpha_n$与基$\alpha_1',\alpha_2',\cdots,\alpha_m'$下的**矩阵**(**表示**)。它是(1)中所示的表示矩阵的转置矩阵。现在，再引入一个形式记号

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)),$$

则(1)式可写成更为简洁的形式

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1', \alpha_2', \cdots, \alpha_m')A. \tag{2}$$

特别，当$V'=V$, $\alpha_i'=\alpha_i$, $i=1, 2, \cdots, n$时，(2)式化为

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)A,$$

其中A为n阶方阵，称A为线性变换σ在基$\alpha_1, \alpha_2, \cdots, \alpha_n$下的**矩阵**(**表示**)。

例1 $R_n[x]$中的微分算子$D(f(x))=f'(x)$在基$1, x, x^2, \cdots, x^n$下的矩阵为如下$n+1$阶上三角阵

$$A = \begin{pmatrix} 0 & 1 & 0 \cdots\cdots 0 \\ 0 & 0 & 2 \cdots\cdots 0 \\ \cdots & \cdots & \cdots\cdots \\ 0 & 0 & 0 \cdots\cdots n \\ 0 & 0 & 0 \cdots\cdots 0 \end{pmatrix}.$$

这是因为，这些基向量在微分算子D之下的像有如下表达式

$$D(1)=0 \quad = 0\cdot 1+0\cdot x+0\cdot x^2+\cdots+0\cdot x^{n-1}+0\cdot x^n,$$
$$D(x)=1 \quad = 1\cdot 1+0\cdot x+0\cdot x^2+\cdots+0\cdot x^{n-1}+0\cdot x^n,$$
$$D(x^2)=2x \quad = 0\cdot 1+2\cdot x+0\cdot x^2+\cdots+0\cdot x^{n-1}+0\cdot x^n,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$D(x^n)=nx^{n-1}=0\cdot 1+0\cdot x+0\cdot x^2+\cdots+n\cdot x^{n-1}+0\cdot x^n.$$

例 2 取定某个复数 $z_0=a+b\sqrt{-1}$，则 $\sigma(z)=z_0\cdot z$ 是复数域 C（它是实数域 R 上的线性空间）上的线性变换，且由

$$\sigma(1)=z_0\cdot 1=a+b\sqrt{-1},$$
$$\sigma(\sqrt{-1})=z_0\cdot\sqrt{-1}=-b+a\sqrt{-1},$$

知 σ 在基 $1, \sqrt{-1}$ 下的矩阵为

$$M=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

例 3 在 $M_n(R)$ 中取基

$$E_{11}, E_{12}, \cdots, E_{1n}; E_{21}, E_{22}, \cdots, E_{2n}; \cdots; E_{n1}, E_{n2}, \cdots, E_{nn}. \tag{3}$$

n 维复行向量全体

$$c_n=\{(z_1, z_2, \cdots, z_n) \mid z_i\in C, i=1, 2, \cdots, n\}$$

可以看成实数域 R 上的 $2n$ 维线性空间，它有基

$$e_1', e_2', \cdots, e_n'; \sqrt{-1}e_1', \sqrt{-1}e_2', \cdots, \sqrt{-1}e_n', \tag{4}$$

这里 e_i' 为第 i 个 n 维单位行向量。考虑 $M_n(R)$ 映到 c_n 内的线性映射

$$\sigma: \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \to (a_{11}, a_{12}, \cdots, a_{1n}).$$

易见，

$$\sigma(E_{ij})=\begin{cases} 1\cdot e_j', & 若 i=1, \\ 0\ , & 若 i\neq 1. \end{cases}$$

故 σ 在基(3)与基(4)下的矩阵为 $2n\times n^2$ 阵

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

因为线性函数是映入基域内的线性映射，它的矩阵应该具有特别简单的形状。

设V是数域K上n维线性空间，f是定义在V上的线性函数。在V中取基$\alpha_1, \alpha_2, \cdots, \alpha_n$，在$K$中取基1，则由

$$f(\alpha_i)=f(\alpha_i)\cdot 1,\ i=1,\ 2,\ \cdots,\ n,$$

知f在基$\alpha_1, \alpha_2, \cdots, \alpha_n$与基1下的矩阵就是$n$维行向量

$$(f(\alpha_1),\ f(\alpha_2),\ \cdots,\ f(\alpha_n)).$$

例 4 K_n上的线性函数

$$f(a_1,\ a_2,\ \cdots,\ a_n)=\sum_{i=1}^{n} a_i$$

在基$e_1', e_2', \cdots, e_n'$与基1下的矩阵就是$(1,\ 1,\ \cdots,\ 1)$。

例 5 $R_n[x]$上的线性函数

$$\sigma:\ p(x)\to\int_0^1 p(x)dx$$

在基$1, x, x^2, \cdots, x^n$与基1下的矩阵是$n+1$维行向量

$$\left(1,\ \frac{1}{2},\ \frac{1}{3},\ \cdots,\ \frac{1}{n+1}\right).$$

利用线性映射σ的矩阵表示可计算出任一向量在σ之下的像向量的坐标。这就是下面的

定理 2.1 设σ是n维线性空间V映入m维线性空间V'的线性映射。σ在V的基$\alpha_1, \alpha_2, \cdots, \alpha_n$与$V'$的基$\alpha_1', \alpha_2', \cdots, \alpha_m'$下的矩阵为$m\times n$阵$A$。任取$\alpha\in V$。若$\alpha$在$\alpha_1, \alpha_2, \cdots, \alpha_n$下的坐标向量为$(k_1, k_2, \cdots, k_n)'$，$\sigma(\alpha)$在$\alpha_1', \alpha_2', \cdots, \alpha_m'$下的坐标向量为$(l_1, l_2, \cdots, l_m)'$，则

$$(l_1,\ l_2,\ \cdots,\ l_m)'=A(k_1,\ k_2,\ \cdots,\ k_n)'. \tag{5}$$

【证明】 由假设条件得

$$\alpha=(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}, \tag{6}$$

$$\sigma(\alpha)=(\alpha_1', \alpha_2', \cdots, \alpha_m')\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{pmatrix} \tag{7}$$

和

$$(\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n))=(\alpha_1', \alpha_2', \cdots, \alpha_m')A.$$

因此，由(6)式得

$$\sigma(\alpha)=(\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n))\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}$$

$$=(\alpha_1', \alpha_2', \cdots, \alpha_m')A\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}. \tag{8}$$

(7)式与(8)式应该相等，再由 $\alpha_1', \alpha_2', \cdots, \alpha_m'$ 线性无关就得到(5)式。 证毕。

特别，若 $V'=V$，$\alpha_i'=\alpha_i$，$i=1, 2, \cdots, n$，则 σ 是 V 上的线性变换。若 σ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为 A，则(5)式就是 α 变到 $\sigma(\alpha)$ 时，在同一个基下的坐标向量之间的变换公式。而第九章定理 3.3 说的是，同一向量在不同的基下的坐标向量变换公式。读者应弄清这两者之间的差别与关系。

二、线性映射与矩阵之间的对应

设 V 与 V' 是数域 K 上两个有限维线性空间。在 V 与 V' 中分别取定基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\alpha_1', \alpha_2', \cdots, \alpha_m'$。第一段的讨论告诉我们，任意给定 V 到 V' 的一个线性映射 σ，必可唯一确定一个 $A\in M_{m,n}(K)$，使得坐标变换公式(5)成立。反之，任意给定一个 $A\in M_{m,n}(K)$，是否必存在 V 到 V' 的线性映射 σ，在相应基下，以此 A 为矩阵表示呢？回答是肯定的。我们先证明以下重要结果。

定理 2.2 设 V 与 V' 分别是数域 K 上的 n 维与 m 维线性空间。设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的基，$\beta_1', \beta_2', \cdots, \beta_n'$ 是 V' 中任意 n 个

向量，则必存在唯一的一个V映入V'的线性映射σ使

$$\sigma(\alpha_i)=\beta_i',\ i=1,2,\cdots,n.$$

【证明】 作如下从V映入V'的映射

$$\sigma:\ \alpha=\sum_{i=1}^n k_i\alpha_i \to \sum_{i=1}^n k_i\beta_i'.$$

因为$\alpha_1,\cdots,\alpha_n$是V的基，所以$k_1,\cdots,k_n$由α唯一确定，而$\beta_1',\cdots,\beta_n'$是取定的，所以σ是单值的。由σ的定义就知它是线性映射，且$\sigma(\alpha_i)=\beta_i',\ i=1,2,\cdots,n.$

今证σ的唯一性。设又存在线性映射τ使

$$\tau(\alpha_i)=\beta_i',\ i=1,2,\cdots,n.$$

则$\sigma(\alpha_i)=\tau(\alpha_i),\ i=1,2,\cdots,n.$ 于是对任一$\alpha=\sum_{i=1}^n k_i\alpha_i\in V$有

$$\sigma(\alpha)=\sum_{i=1}^n k_i\sigma(\alpha_i)=\sum_{i=1}^n k_i\tau(\alpha_i)=\tau(\alpha).$$

这就说明$\sigma=\tau$。 证毕。

定理 2.3 设V与V'分别是数域K上的n维与m维线性空间，在V与V'中分别取定基$\alpha_1,\alpha_2,\cdots,\alpha_n$与$\alpha_1',\alpha_2',\cdots,\alpha_m'$。任给$A\in M_{m\times n}(K)$，则必存在唯一的$V$映入$V'$的线性映射$\sigma$，它在基$\alpha_1,\alpha_2,\cdots,\alpha_n$与基$\alpha_1',\alpha_2',\cdots,\alpha_m'$下的矩阵就是这个$A$。

【证明】 设

$$A=\begin{pmatrix} a_{11} & a_{21}\cdots a_{n1} \\ a_{12} & a_{22}\cdots a_{n2} \\ \cdots & \cdots \\ a_{1m} & a_{2m}\cdots a_{nm}\end{pmatrix}_{m\times n}$$

(注意A中元素的排列)。造V'中如下n个向量

$$\beta_i'=\sum_{j=1}^m a_{ij}\alpha_j',\ i=1,2,\cdots,n.$$

由定理2.2，存在V映入V'的唯一的线性映射σ使

$$\sigma(\alpha_i)=\beta_i'=\sum_{j=1}^m a_{ij}\alpha_j',\ i=1,2,\cdots,n.$$

这n个等式恰好说明σ在基$\alpha_1,\alpha_2,\cdots,\alpha_n$与基$\alpha_1',\alpha_2',\cdots,\alpha_m'$下的矩阵就是这个给定的$A$。 证毕。

定理 2.4 设 V 与 V' 分别是数域 K 上的 n 维与 m 维线性空间，则在取定一对基 $\alpha_1, \alpha_2, \cdots, \alpha_n$（属于 V）与 $\alpha'_1, \alpha'_2, \cdots, \alpha'_m$（属于 V'）的前提下，V 映入 V' 的线性映射集合 $L(V, V')$ 与 $m\times n$ 矩阵集合 $M_{m,n}(K)$ 之间存在一一对应。

【证明】 对任一 $\sigma\in L(V, V')$，作对应

$$\eta:\ \sigma\to A,$$

这里 A 取作 σ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与基 $\alpha'_1, \alpha'_2, \cdots, \alpha'_m$ 下的矩阵。显然，$A\in M_{m,n}(K)$。因为这个 A 由 σ 唯一确定，所以 η 是单值映射。定理 2.3 保证了 η 是单射和满射，所以 η 是一一对应。证毕。

如此建立起来的一一对应 η，有一些重要性质，将在 § 5 中介绍。

三、线性映射在不同基下的矩阵之间的关系式

设 σ 是 n 维线性空间 V 映入 m 维线性空间 V' 的线性映射，它在 V 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 V' 的基 $\alpha'_1, \alpha'_2, \cdots, \alpha'_m$ 下的矩阵为 A，在 V 的基 $\beta_1, \beta_2, \cdots, \beta_n$ 与 V' 的基 $\beta'_1, \beta'_2, \cdots, \beta'_m$ 下的矩阵为 B。问 A 与 B 之间有什么关系？我们将证明它们是相抵的。

定理 2.5 如果 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 到 $\beta_1, \beta_2, \cdots, \beta_n$ 的过渡矩阵是 n 阶阵 P，$\alpha'_1, \alpha'_2, \cdots, \alpha'_m$ 到 $\beta'_1, \beta'_2, \cdots, \beta'_m$ 的过渡矩阵是 m 阶阵 Q，则 $B=Q^{-1}AP$。

【证明】 由假设条件可得以下诸等式

$$(\beta_1, \beta_2, \cdots, \beta_n)=(\alpha_1, \alpha_2, \cdots, \alpha_n)P, \tag{9}$$

$$(\beta'_1, \beta'_2, \cdots, \beta'_m)=(\alpha'_1, \alpha'_2, \cdots, \alpha'_m)Q, \tag{10}$$

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n)=(\alpha'_1, \alpha'_2, \cdots, \alpha'_m)A, \tag{11}$$

$$\sigma(\beta_1, \beta_2, \cdots, \beta_n)=(\beta'_1, \beta'_2, \cdots, \beta'_m)B. \tag{12}$$

把(9)、(10)两式代入(12)式，并注意到形式记号的含义得

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n)P=(\alpha'_1, \alpha'_2, \cdots, \alpha'_m)QB.$$

用(11)式代入得

$$(\alpha'_1, \alpha'_2, \cdots, \alpha'_m)AP=(\alpha'_1, \alpha'_2, \cdots, \alpha'_m)QB.$$

再据 $\alpha'_1, \alpha'_2, \cdots, \alpha'_m$ 线性无关即得 $AP=QB$。因为两个基之间的

过渡矩阵Q必是可逆阵，所以$B=Q^{-1}AP$。　　　　证毕。

特别，在定理2.5中取$V'=V$，$\alpha_i'=\alpha_i$，$\beta_i'=\beta_i$，$i=1,2,\cdots,n$，则$Q=P$，$B=P^{-1}AP$。于是得

定理2.6　线性空间中任一线性变换在不同基下的矩阵是相似的。

定理2.6实际上给我们提供了求n阶阵A的相似标准形的另一途径(与第五章、第六章中所说的不同)。由定理2.3，可把这个A看作n维线性空间V中某一线性变换σ在某基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的矩阵，则$\sigma$在另一个基$\beta_1,\beta_2,\cdots,\beta_n$下的矩阵就是$B=P^{-1}AP$，这里$P$是$\alpha_1,\alpha_2,\cdots,\alpha_n$到$\beta_1,\beta_2,\cdots,\beta_n$的过渡矩阵，所以问题变为：如何找适当的基$\beta_1,\beta_2,\cdots,\beta_n$使对应的$B$有最简单的形状。所以求矩阵相似标准形的问题也可用线性空间中的线性变换的理论解决。

习　题

1．求出下列线性空间中的线性映射(或线性变换)在所给基下的矩阵。

(i) $M_n(R)$映入$R^{(n)}$内的线性映射

$$A\to A\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix},$$

对应的基分别取E_{ij}，$i,j=1,2,\cdots,n$和e_j，$j=1,2,\cdots,n$；

(ii) n维线性空间V的投影(射影)变换

$$\sum_{i=1}^{n}k_i\alpha_i\to\sum_{i=1}^{m}k_i\alpha_i,$$

这里$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$的基，$m$是小于$n$的取定的自然数；

(iii) 复数域$C=R(\sqrt{-1})$中的共轭变换$z\to\bar{z}$，取基为$1,\sqrt{-1}$；

(iv) $M_n(K)$映入K的线性函数

$$A=(a_{ij})_{n\times n}\to\sum_{i,j=1}^{n}a_{ij},$$

基分别取E_{ij}，$i,j=1,2,\cdots,n$和1；

(v) $R_n[x]$的线性变换：$f(x)\to f(x+1)$，取基为$1,x,\cdots,x^n$；

(vi) $R_n[x]$的线性变换：$f(x)\to f(x+1)-f(x)$，取基为$1,x,\cdots,x^n$；

(vii) $R_n[x]$的线性变换：$f(x)\to f'(x)$，取基为

$$1,\ \frac{x}{1!},\ \frac{x^2}{2!},\ \cdots,\ \frac{x^n}{n!};$$

(viii) R_2 的线性变换

$$(1, 0)\to(0, 0),\ (0, -1)\to(3, 2),$$

基取为$(1, 1)$, $(0, 1)$。

2. 求 n 维线性空间V的零变换 0^* 与单位变换 1^* 在V的任何基下的矩阵。

3. 求数域K上 n 维线性空间V中的线性变换 $\alpha\to k\alpha$ 在V的任何基下的矩阵，这里 k 是K中取定的一个数。

4. 设σ是数域K上n维线性空间V中一个线性变换。如果σ在V的任何基下的矩阵都相同。求证：必存在 $k\in K$ 使 $\sigma(\alpha)=k\alpha$, $\forall\alpha\in V$。

5. 在 K_3 中取定

$$\alpha_1=(-1, 0, 2),\ \alpha_2=(0, 1, 1),\ \alpha_3=(3, -1, 0).$$

作如下线性变换 σ：

$$\sigma(\alpha_1)=(-5, 0, 3),\ \sigma(\alpha_2)=(0, -1, 6),\ \sigma(\alpha_3)=(-5, -1, 9).$$

求 σ 在基 $e_1'=(1, 0, 0)$, $e_2'=(0, 1, 0)$, $e_3'=(0, 0, 1)$下的矩阵 A 和在 α_1, α_2, α_3 下的矩阵 B。

§3 线性映射的象空间与核空间

在第九章§4末，对任一 $A\in M_{m\times n}(R)$，定义了A的象空间W与核空间N如下：

$$W=\{Ax\,|\,x\in R^{(n)}\},$$
$$\boldsymbol{N}=\{x\,|\,Ax=0,\ x\in R^{(n)}\}.$$

且证明了W是 $R^{(m)}$的 r 维子空间，N 是 $R^{(n)}$的 $n-r$ 维子空间，这里 r 是A的秩。于是有

$$\mathrm{d}(W)+\mathrm{d}(\boldsymbol{N})=n=\mathrm{d}(R^{(n)}).\tag{1}$$

在§2 中已说过，由这个A又可确定$R^{(n)}$映入$R^{(m)}$的一个线性映射(我们也用A表示)：

$$A:\ x\to Ax,\ \forall x\in R^{(n)}.$$

所以W和N也可认为是 $R^{(n)}$映入 $R^{(m)}$的线性映射A的象空间和核

空间.

现在,把象空间与核空间的概念推广到两个线性空间之间的任一线性映射上去,且对有限维空间来说,仍有维数关系式(1).

设 V 和 V' 是数域 K 上任意两个线性空间, σ 是 V 映入 V' 的线性映射. 记

$$\sigma(V)=\{\sigma(\alpha)\mid\alpha\in V\},$$
$$\sigma^{-1}(\theta')=\{\alpha\mid\sigma(\alpha)=\theta',\alpha\in V\},$$

其中 θ' 是 V' 的零向量. 注意, 这里 $\sigma^{-1}(\theta')$是个完整的记号, 并不是说 σ 是可逆映射.

命题1 $\sigma(V)$是 V' 的子空间. $\sigma^{-1}(\theta')$是 V 的子空间.

【证明】 由 $\sigma(\theta)=\theta'$ 知 $\sigma(V)$ 和 $\sigma^{-1}(\theta')$ 都不是空集. 任取 $\sigma(\alpha),\sigma(\beta)\in\sigma(V)$, $k,l\in K$ 都有

$$k\sigma(\alpha)+l\sigma(\beta)=\sigma(k\alpha+l\beta)\in\sigma(V),$$

所以 $\sigma(V)$是 V' 的子空间. 任取 $\alpha,\beta\in\sigma^{-1}(\theta')$, $k,l\in K$, 都有

$$\sigma(k\alpha+l\beta)=k\sigma(\alpha)+l\sigma(\beta)=\theta',$$

所以 $\sigma^{-1}(\theta')$是 V 的子空间. 证毕.

定义 称 $\sigma(V)$为 σ 的**象空间**或**值域**, $\sigma(V)$ 的维数称为 σ 的**秩**. 称 $\sigma^{-1}(\theta')$为 σ 的**核空间**或**核**.

特别,若 σ 是线性空间 V 的线性变换, 则 $\sigma(V)$和 $\sigma^{-1}(\theta)$都是 V 的子空间.

命题2 V 映入 V' 的线性映射 σ 是单射的充分必要条件是 $\sigma^{-1}(\theta')=\theta$.

【证明】 若 σ 是单射. 任取 $x\in\sigma^{-1}(\theta')$, 则 $\sigma(x)=\theta'$, 但 $\sigma(\theta)=\theta'$, 所以由 σ 是单射知 $x=\theta$.

若 $\sigma^{-1}(\theta')=\theta$. 如果 $\sigma(x)=\sigma(y)$,则 $\sigma(x-y)=\theta'$. $x-y\in\sigma^{-1}(\theta')$, 所以 $x=y$. 这说明 σ 是单射. 证毕.

例1 取定某个非零复数 z_0, 它确定复数域 C 的线性变换 $\sigma(z)=z_0\cdot z$. 因为对任一 $z\in C$, 必可写成

$$z=z_0(z_0^{-1}z)=\sigma(z_0^{-1}z),$$

所以 σ 是满射; $\sigma(C)=C$. 这说明 σ 的象空间就是 C 本身. 因为

复数的消去律成立，即由 $z_0z=z_0z'$ 必可推出 $z=z'$，所以 σ 是单射，$\sigma^{-1}(0)=0$。这说明 σ 的核空间是零空间。

例 2 R_n 上的线性函数

$$f:\ (a_1,a_2,\cdots,a_n)\to\sum_{i=1}^{n}a_i$$

显然是一个满射，所以 f 的象空间 $f(R_n)=R$，而 f 的核空间为

$$f^{-1}(0)=\{(a_1,a_2,\cdots,a_n)\,|\,a_i\in R,\ \sum_{i=1}^{n}a_i=0\}.$$

命题 3 设 σ 是 n 维线性空间 V 映入 m 维线性空间 V' 的线性映射，σ 在 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 V' 的基 $\alpha'_1,\alpha'_2,\cdots,\alpha'_m$ 下的矩阵是 $m\times n$ 阵 A，则

$$\sigma(V)=L(\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)),\tag{2}$$

且 $\mathrm{d}(\sigma(V))=\mathrm{r}(A)$，这就是说，$\sigma$ 的秩等于 A 的秩。

【证明】 因为 $\sigma(\alpha_i)\in\sigma(V)$ 而 $\sigma(V)$ 是 V' 的子空间，所以

$$L(\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n))\subseteq\sigma(V).$$

反之，任取 $\alpha=\sum_{i=1}^{n}k_i\alpha_i\in V$，有

$$\sigma(\alpha)=\sum_{i=1}^{n}k_i\sigma(\alpha_i)\in L(\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)),$$

所以(2)式成立。

据第九章定理 4.3，$\sigma(V)$ 的维数等于 $\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)$ 中线性无关向量的最大个数。再据第九章 § 6 的习题 3 知道，它就是 A 的秩。 证毕。

因为线性变换 σ 的秩，也就是象空间 $\sigma(V)$ 的维数，是由 σ 本身所确定的，与基的选择无关，所以命题 3 进一步说明了，不管基如何选择，所得的矩阵 A 可以不同，但它们的秩都等于 $\mathrm{d}(\sigma(V))$。

下面我们来推导类似于(1)的维数关系式。

定理 3.1 设 σ 是 n 维线性空间 V 映入 m 维线性空间 V' 的线性映射，θ' 是 V' 的零向量，则

$$\mathrm{d}(\sigma(V))+\mathrm{d}(\sigma^{-1}(\theta'))=n=\mathrm{d}(V)\tag{3}$$

（这个等式不牵涉到 V' 的维数 m）。

【证明】 分两种情形：

(i) 如果 σ 是单射，此时，$\mathrm{d}(\sigma^{-1}(\theta'))=0$。任取 V 的基 α_1，α_2，…，α_n。据§1的命题2知 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，…，$\sigma(\alpha_n)$ 线性无关，所以由

$$\sigma(V)=L(\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n))$$

得 $\mathrm{d}(\sigma(V))=n$。此时(3)式成立。

(ii) 如果 σ 不是单射。此时，由命题2，$\mathrm{d}(\sigma^{-1}(\theta'))=s>0$。在 V 的 s 维子空间 $\sigma^{-1}(\theta')$ 中可取某基 $\alpha_1,\alpha_2,\cdots,\alpha_s$，并可把它扩充成 V 的基 $\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$。注意到，对 $i=1,2,\cdots,s$，有 $\sigma(\alpha_i)=\theta'$，所以对 $\alpha=\sum\limits_{i=1}^{n}k_i\alpha_i$ 有

$$\sigma(\alpha)=\sum_{i=s+1}^{n}k_i\sigma(\alpha_i),$$

所以

$$\sigma(V)=L(\sigma(\alpha_{s+1}),\sigma(\alpha_{s+2}),\cdots,\sigma(\alpha_n)).\tag{4}$$

再证 $\sigma(\alpha_{s+1}),\sigma(\alpha_{s+2}),\cdots,\sigma(\alpha_n)$ 线性无关。设

$$\sum_{i=s+1}^{n}l_i\sigma(\alpha_i)=\theta',$$

这说明

$$\sum_{i=s+1}^{n}l_i\alpha_i\in\sigma^{-1}(\theta'),$$

所以

$$\sum_{i=s+1}^{n}l_i\alpha_i=\sum_{i=1}^{s}k_i\alpha_i,$$

但 $\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$ 线性无关，所以 $l_i=0$，$i=s+1$，…，n，$\sigma(\alpha_{s+1}),\sigma(\alpha_{s+2}),\cdots,\sigma(\alpha_n)$ 线性无关，于是据(4)式即得

$$\mathrm{d}(\sigma(V))=n-s=n-\mathrm{d}(\sigma^{-1}(\theta')).$$

这就是所要证明的(3)式。 证毕。

顺便指出，应用第九章§4末的结论和线性空间的同构可直接得到定理3.1。

特别，取 $V'=V$，则定理3.1是说，对于 V 的任一线性变换 σ

有

$$\mathrm{d}(\sigma(V))+\mathrm{d}(\sigma^{-1}(\theta))=\mathrm{d}(V)。\tag{5}$$

要注意的是,不能仅仅根据这个维数等式(5)就说

$$V=\sigma(V)\oplus\sigma^{-1}(\theta)。$$

事实上,这未必正确。例如,$R_n[x]$中的微分算子

$$D:\ p(x)\to p'(x)$$

的象空间是 $R_{n-1}[x]$,核空间是 R,两者维数之和为 n,但 $R_{n-1}[x]\cap R=R$,所以 $R_n[x]\neq R_{n-1}[x]\oplus R$。

一般说来,某映射 σ 是不是单射与是不是满射,这两者之间并无必然的联系(见 § 1 的有关例子)。当然,如果 σ 是某两个元素个数相同的有限集合之间的映射,则 σ 是单射当且仅当 σ 是满射。现在,由定理 3.1 可以推出,这个事实竟然对有限维线性空间(它有无限多个元素)中的线性变换也是正确的,这就是

推论 设 σ 是有限维线性空间 V 中的线性变换,则 σ 是单射当且仅当 σ 是满射。

【证明】 若 σ 是单射。据定理 3.1 的证明的第一部分知 $\mathrm{d}(\sigma(V))=\mathrm{d}(V)$,故 $\sigma(V)=V$,σ 是满射。

设 σ 是满射,则 $\mathrm{d}(\sigma(V))=\mathrm{d}(V)$。如果 σ 不是单射,则由定理 3.1 的证明的第二部分知必有 $\mathrm{d}(\sigma(V))<\mathrm{d}(V)$,这是一个矛盾。所以 σ 必是单射。　证毕。

因此,对有限维线性空间中的线性变换来说,单射、满射和双射都是同一回事。值得提醒读者注意的是:由定理 3.1 的结论可知,对两个不同的有限维线性空间之间的线性映射来说,这个结论未必成立。事实上,在 §1 中已举出很多例子,其中 σ 是满射,但不是单射,或者是单射,但不是满射。

习　题

1. 求下列基域 K 上的线性空间 V 映入 V' 的线性映射 σ 的象空间 W 与核空间 N,且当 V 与 V' 都是有限维时,求出 W 和 N 的维数。

(i) $V=R_n[x]$, $V'=R$, $K=R$,

$$\sigma: \ p(x) \to \int_0^1 p(x)\mathrm{d}x;$$

(ii) $V = M_n(R)$, $V' = C_n$, $K = R$,

$$\sigma: \ (a_{ij})_{n\times n} \to (a_{11}, a_{12}, \cdots, a_{1n});$$

(iii) $V = M_3(R)$, $V' = M_{2,3}(R)$, $K = R$, $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$,

$$\sigma: \ X \to AX;$$

(iv) $V = M_n(R)$, $V' = R$, $K = R$,

$$\sigma: \ A \to \mathrm{tr}(A);$$

(v) $V = M_n(R)$, $V' = R$, $K = R$,

$$\sigma: \ (a_{ij})_{n\times n} \to \sum_{i,j=1}^{n} a_{ij};$$

(vi) $V = C$, $K = R$, $\sigma: \ z \to \bar{z}$;

(vii) $V = R_n[x]$, $K = R$, $\sigma: \ f(x) \to f(x+1)$;

(viii) $V = R_n[x]$, $K = R$, $\sigma: f(x) \to f(x+1) - f(x)$;

(ix) V 的零变换与单位变换;

(x) $V = R[a, \infty)$, $K = R$,

$$\sigma: \ f(x) \to \int_a^x f(t)\mathrm{d}t.$$

2. 设σ是有限维线性空间V的线性变换,W是V的子空间,求证

$$\mathrm{d}(\sigma(W)) + \mathrm{d}(\sigma^{-1}(0) \cap W) = \mathrm{d}(W),$$

它是(5)式的推广。

§4 不变子空间、线性变换的特征值与特征向量

一、不变子空间、诱导线性变换

设W是K上线性空间V的某个子空间,σ是V的线性变换。一般说来,对$\alpha \in W$,未必有$\sigma(\alpha) \in W$。但也确实存在对所有$\alpha \in W$都有$\sigma(\alpha) \in W$,即$\sigma(W) \subseteq W$的情形。

例 1 $R[x]$的微分算子$D: p(x) \to p'(x)$。对$R[x]$的$n+1$维子空间$R_n[x]$有

$$D(R_n[x]) = R_{n-1}[x] \subseteq R_n[x].$$

例 2 设 $\Omega(-\infty, \infty)$是定义在$(-\infty, \infty)$上的任意次可微实函数空间．则对微分算子 $D: f(x) \to f'(x)$有 $D(R[x]) = R[x]$，而$R[x]$是 $\Omega(-\infty, \infty)$的子空间。

例 3 设 $V = L(\alpha_1, \alpha_2, \cdots, \alpha_n)$是 n 维线性空间，$\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的基。$W = L(\alpha_1, \alpha_2, \cdots, \alpha_m)$是 V 的子空间，$m<n$，则投影(射影)变换

$$P: \sum_{i=1}^{n} k_i\alpha_i \to \sum_{i=1}^{m} k_i\alpha_i$$

满足 $P(W) = W$。

定义 设V是任意线性空间(有限维或无限维)，W是V的子空间，σ是V的线性变换。如果$\sigma(W)\subseteq W$，则称W是V的关于σ的**不变子空间**，也可简称为σ-**不变子空间**。

易见，V的任一子空间W必是零变换 0^* 和恒等变换 1^* 的不变子空间。零空间θ和V本身都是V中任一线性变换σ的不变子空间，称θ和V为**平凡不变子空间**。

例 4 V的任何子空间W都是**纯量线性变换**

$$\sigma: \alpha \to k\alpha$$

的不变子空间，这里k是基域K中取定的某个非零数。

命题 1 设σ是线性空间V的任一线性变换，则$\sigma(V)$与$\sigma^{-1}(\theta)$都是σ的不变子空间。

【证明】 任取 $\beta\in\sigma(V)\subseteq V$，则 $\sigma(\beta)\in\sigma(V)$，所以 $\sigma(\sigma(V))\subseteq\sigma(V)$。任取 $\alpha\in\sigma^{-1}(\theta)$，则 $\sigma(\alpha)=\theta$，所以 $\sigma(\sigma^{-1}(\theta))=\theta\subseteq\sigma^{-1}(\theta)$。

证毕。

设W是线性空间V的关于线性变换σ的不变子空间。既然，对任一$\alpha\in W$，必有$\sigma(\alpha)\in W$，所以σ也可看作是线性空间W中的线性变换。我们以σ_W记之，即

$$\sigma_W(\alpha)=\sigma(\alpha),\ \forall\alpha\in W. \tag{1}$$

定义 由(1)式确定的σ_W称为σ在其不变子空间W上的**诱导(或导出)线性变换**。

注意：σ_W在W上与σ一致，但在W以外，则认为σ_W是没有定

义的。

例 5 σ 在 $\sigma^{-1}(\theta)$上的诱导变换是零变换。即

$$\sigma_{\sigma^{-1}(\theta)} = 0^*.$$

引进不变子空间的目的之一是，可借助于不变子空间的基，求出线性变换的矩阵在相似之下的简化形状。

设 σ 是 K 上 n 维线性空间 V 的线性变换，W是 σ 的不变于空间。任取W的基 $\alpha_1,\alpha_2,\cdots,\alpha_s$，并把它扩充成 V 的基 $\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$。则由 $\sigma(W)\subseteq W$ 立得

$$\begin{aligned}
&\sigma(\alpha_1) = a_{11}\alpha_1 + \cdots + a_{1s}\alpha_s\\
&\sigma(\alpha_2) = a_{21}\alpha_1 + \cdots + a_{2s}\alpha_s\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\sigma(\alpha_s) = a_{s1}\alpha_1 + \cdots + a_{ss}\alpha_s \qquad\qquad (2)\\
&\sigma(\alpha_{s+1}) = a_{s+1,1}\alpha_1 + \cdots + a_{s+1,s}\alpha_s + a_{s+1,s+1}\alpha_{s+1} + \cdots + a_{s+1,n}\alpha_n\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\sigma(\alpha_n) = a_{n1}\alpha_1 + \cdots + a_{ns}\alpha_s + a_{n,s+1}\alpha_{s+1} + \cdots + a_{nn}\alpha_n
\end{aligned}$$

所以 σ 在此基下的矩阵是形如

$$\boldsymbol{A} = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix} \qquad (3)$$

的分块上三角阵，其中

$$\boldsymbol{A}_1 = \begin{pmatrix} a_{11} & a_{21}\cdots a_{s1} \\ a_{12} & a_{22}\cdots a_{s2} \\ \cdots\cdots & \cdots\cdots \\ a_{1s} & a_{2s}\cdots a_{ss} \end{pmatrix},\quad A_2 = \begin{pmatrix} a_{s+1,1}\cdots\cdots a_{n1} \\ a_{s+1,2}\cdots\cdots a_{n2} \\ \cdots\cdots\cdots\cdots \\ a_{s+1,s}\cdots\cdots a_{ns} \end{pmatrix},$$

$$O = O_{(n-s)\times s},\qquad \boldsymbol{A}_3 = \begin{pmatrix} a_{s+1,s+1}\cdots\cdots a_{n,s+1} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s+1,n} \quad \cdots\cdots a_{nn} \end{pmatrix},$$

由(2)式可知，σ_W 在W的基 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 下的矩阵就是 $\boldsymbol{A}_1$。

特别，若 $s=1$，即W是 σ 的一维不变子空间，则(3)式为

$$A=\begin{pmatrix} a & \beta \\ 0 & A_3 \end{pmatrix},$$

这里 $a\in K$，β 为 K 上 $n-1$ 维行向量，$A_3\in M_{n-1}(K)$。由前几章的讨论可知，具这种形状的矩阵是特别重要的。

由上述推理可知，利用一个不变子空间 W，可使 σ 的表示矩阵 A 中产生一个零块。如果 $V=W\oplus U$，而 W 和 U 都是 σ 的不变子空间，则可分别取 W 的基 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 与 U 的基 $\alpha_{s+1},\alpha_{s+2},\cdots,\alpha_n$ 合成 V 的基，于是 σ 在此基下的矩阵必为分块对角阵：

$$A=\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

这里 A_1 和 A_2 分别是 s 阶和 $n-s$ 阶方阵。进一步，如果能把 V 分解为 σ 的若干个不变子空间的直和，则在相应的基下，σ 的矩阵就有较简单的分块对角形状。直和分解式分得越细，即不变子空间直和因子的维数越小，则矩阵的形状就越简单，这就是我们要讨论有限维线性空间的关于线性变换的不变子空间的主要目的。

二、线性变换的特征值与特征向量

上面已经提到，有限维线性空间的关于线性变换 σ 的一维不变子空间是重要的。但是，并非对任一 σ 都存在一维不变子空间，即使在无限维线性空间中也是如此。

定理 4.1 设 σ 是数域 K 上任意线性空间 V 的线性变换，则 σ 有一维不变子空间的充分必要条件是，存在非零向量 $\alpha\in V$ 和 $\lambda\in K$ 使

$$\sigma(\alpha)=\lambda\alpha. \tag{4}$$

此时，$W=L(\alpha)$ 就是 V 的关于 σ 的一维不变子空间。

【证明】 设 W 是 σ 的一维不变子空间，则 $W=L(\alpha)$。再由 $\sigma(W)\subseteq W$ 知存在 $\lambda\in K$ 使 $\sigma(\alpha)=\lambda\alpha$。反之，设 $\sigma(\alpha)=\lambda\alpha$。令 $W=L(\alpha)$。任取 $\beta\in W$，则 $\beta=k\alpha$，$k\in K$，而 $\sigma(\beta)=k\sigma(\alpha)=k\lambda\alpha\in L(\alpha)=W$，所以 $W=L(\alpha)$ 就是 σ 的一维不变子空间。 证毕。

这说明，使(4)式成立的向量 α 与数 λ 有特殊的重要性。为此

引进

定义 称满足(4)式的数$\lambda\in K$为线性变换σ的**特征值**，称对应的非零向量α为σ的属于λ的**特征向量**。

例 6 $R[x]$中关于微分算子D有且仅有一个一维不变子空间，就是R，且R就是D的核空间。

事实上，由$D(p(x))=p'(x)$易知R是D的不变子空间，维数为1，且R就是D的核。若W是$R[x]$的任意一个一维D-不变子空间，则$W=L(p(x))$，且存在$\lambda\in R$使$p'(x)=\lambda\cdot p(x)$，比较两端的次数即得$p(x)=k\in R$，$\lambda=0$，所以$W=L(k)=R$。同时证得$\lambda=0$是D的唯一的特征值。任一非零实数k都是属于$\lambda=0$的特征向量，因而有无限多个。

例 7 设$\Omega[a,b]$是定义在$[a, b]$上的任意次可微实函数所成的实线性空间。确定$\Omega[a,b]$中关于微分算子D的所有一维不变子空间。

首先，同例6一样，可证$\lambda=0$是D的一个特征值，且任一非零实数k所生成的$L(k)$都是属于$\lambda=0$的一维不变子空间，也就是基域R。

其次，任取$0\neq\lambda\in R$。如果λ是D的一个特征值，则必存在$0\neq f(x)\in\Omega[a, b]$使$f'(x)=\lambda f(x)$，通过积分可证必有$f(x)=ke^{\lambda x}$，这里$k$是任意非零实数。由此可见，任一非零实数$\lambda$都是$D$的特征值，而对任一非零实数$k$，函数$f(x)=ke^{\lambda x}$都是$D$的属于$\lambda$的特征向量。$D$的一维不变子空间必为$L(e^{\lambda x})$。

综上两点所述可知，$L(e^{\lambda x})$是D的所有的一维不变子空间，其中λ是任意实数。任一实数λ都是D的特征值。对任一非零实数k，$ke^{\lambda x}$都是属于λ的特征向量。

例 8 实线性空间$\Omega[a,b]$中没有关于线性变换

$$I(f(x))=\int_a^x f(t)\mathrm{d}t$$

的一维不变子空间。

【证明】 设存在$\lambda\in R$与$0\neq f(x)\in R[a,b]$使

$$\int_a^x f(t)dt = \lambda f(x), \tag{5}$$

则 $\lambda \neq 0$(否则,必有 $f(x)=0$),在(5)式中取 $x=a$,得

$$\lambda \cdot f(a) = 0, f(a) = 0,$$

将(5)式两边求导数得 $f(x)=\lambda \cdot f'(x)$,于是必有

$$f(x) = ke^{\frac{1}{\lambda}x}, \ k \text{ 是非零实数}.$$

但另一方面,由 $f(a)=0$ 又得 $ke^{\frac{1}{\lambda}a}=0$,这为矛盾。　　　　证毕。

例 9　复数域 C 是有理数域 Q 上无限维线性空间。求 C 中共轭线性变换 $\sigma: z \to \bar{z}$ 的所有特征值与特征向量。并求出 C 的关于 σ 的所有一维不变子空间。

设 $\lambda \in Q$, $0 \neq z = a + b\sqrt{-1} \in C$ 满足 $\bar{z} = \lambda z$,即

$$a - b\sqrt{-1} = \lambda(a + b\sqrt{-1}),$$

则由 $z \neq 0$ 知 $\lambda \neq 0$,于是只有两种可能性:

(i) $\lambda = 1, b = 0$。此时,任意非零实数 a 都是属于特征值 $\lambda = 1$ 的特征向量。$L(a)$都是 σ 的不变子空间。

(ii) $\lambda = -1, a = 0$。此时,任意非零纯虚数 $b\sqrt{-1}$ 都是属于特征值 $\lambda = -1$ 的特征向量。$L(b\sqrt{-1})$都是 σ 的不变子空间。

现在,我们进一步分析关系式 $\sigma(\alpha) = \lambda\alpha$, $\alpha \in V$, $\lambda \in K$, σ 是 V 的线性变换。如果 α 与 β 同时是 V 中属于特征值 λ 的特征向量:$\sigma(\alpha) = \lambda\alpha, \sigma(\beta) = \lambda\beta$,则对任何 $k, l \in K$,也有

$$\sigma(k\alpha + l\beta) = k\sigma(\alpha) + l\sigma(\beta) = \lambda(k\alpha + l\beta),$$

这说明 $k\alpha + l\beta$ 都是属于 λ 的特征向量。零向量虽然不认为是特征向量,但也满足 $\sigma(\alpha) = \lambda\alpha$,于是证得集合

$$V_\lambda = \{\alpha \mid \alpha \in V \text{ 满足 } \sigma(\alpha) = \lambda\alpha\}$$

是 V 的子空间,这里 λ 是 σ 的取定的一个特征值。易见,V_λ是 σ 的不变子空间。我们称 V_λ 为(**属于特征值 λ 的**)**特征子空间**。当然,它的维数未必是 1。

在例 9 中,有两个特征子空间:

$$V_1 = \{z \mid z \in C, \bar{z} = z\} = R,$$

$$V_{-1}=\{z|z\in C,\bar{z}=-z\}=\{b\sqrt{-1}|b\in R\}.$$

它们都是有理数域Q上的无限维空间。再由

$$V_1\cap V_{-1}=0, C=V_1+V_{-1},$$

可得

$$C=V_1\oplus V_{-1}.$$

若σ是无限维线性空间V的线性变换，则要找出σ的一维不变子空间，有时是比较困难的。然而，当V是有限维时，却比较容易找出。我们有

定理 4.2 设V是K上n维线性空间，σ是V的线性变换，σ在V的某一基下的矩阵为A，则某个数$\lambda_0\in K$是σ的特征值的充分必要条件是，λ_0是矩阵A的特征值，即是A的特征多项式

$$f(\lambda)=|\lambda I_n-A|$$

的根。此时，σ的属于λ_0的特征向量就是A的属于λ_0的特征向量。

【证明】 设σ在V的基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的矩阵为$A$。$V$中任一向量$\alpha$在此基下的坐标向量为$x=(x_1,x_2,\cdots,x_n)'$。由定理 2.1 知$\sigma(\alpha)$在此基下的坐标向量就是$Ax$。对任一$\lambda\in K$，$\lambda\alpha$在此基下的坐标向量为$\lambda x$，于是据坐标向量的唯一性，$\sigma(\alpha)=\lambda_0\alpha$成立当且仅当$Ax=\lambda_0 x$成立，即$\lambda_0$是$f(\lambda)$的根。由此，根据特征值和特征向量的有关定义即证得定理的结论。 证毕。

现在，我们指出一个重要事实：设σ是K上n维线性空间V的线性变换。由定理 2.6 知σ在V的两个不同的基下的矩阵是相似的，而相似矩阵有相同的特征多项式$f(\lambda)$，所以可把这个$f(\lambda)$称为这个σ**的特征多项式**。于是，定理 4.2 也可叙述为

定理 4.2′ 设σ是K上n维线性空间V的线性变换，则$\lambda_0\in K$是σ的特征值的充分必要条件是，λ_0是σ的特征多项式在K中的根。

因为K上的多项式$f(\lambda)$在K中未必有根，所以线性变换σ的特征值未必存在。如果K是复数域，则由代数基本定理知$f(\lambda)$在K中必有根，故有

推论 有限维复线性空间中的任一线性变换必有一维不变子

空间。

习　题

1. 设V_1与V_2是线性空间V中线性变换σ的不变子空间，求证：V_1+V_2，$V_1\cap V_2$也是σ的不变子空间。

2. 复数域C是二维实线性空间。取定$z_0=a+b\sqrt{-1}\in C$。求证：线性变换$\sigma: z\to z_0\cdot z$有非平凡的不变子空间的充分必要条件是$b=0$。

3. 设$\alpha_1, \alpha_2, \cdots, \alpha_n$是$K$上线性空间$V$的基。取定

$$A=\begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}\in M_n(K),$$

其中$A_1\in M_r(K)$，$r<n$。求证：$L(\alpha_1, \alpha_2, \cdots, \alpha_r)$是$V$中线性变换

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n)=(\alpha_1, \alpha_2, \cdots, \alpha_n)A$$

的不变子空间。

4. 设有限维线性空间$V=V_1\oplus V_2\oplus\cdots\oplus V_s$，$\sigma$是$V$的线性变换，$V_i$都是$\sigma$的不变子空间。求证：必可选择$V$的基，使$\sigma$在此基下的矩阵为分块对角阵。

5. 设W是有限维线性空间V的子空间，且是V中某个线性变换σ的不变子空间。如果σ是双射，证明W也是σ^{-1}的不变子空间。

§5　线性变换的运算

一、线性映射的加法与数乘

定义　设V与V'是数域K上的线性空间，σ与τ是V映入V'的任意两个映射，对$\alpha\in V$，$k\in K$，定义

$$(\sigma+\tau)(\alpha)=\sigma(\alpha)+\tau(\alpha),$$

$$(k\sigma)(\alpha)=(\sigma k)(\alpha)=k\sigma(\alpha),$$

则$\sigma+\tau$与$k\sigma=\sigma k$都是V映入V'的映射，分别称为**σ与τ的和**，**k与σ的数乘**。

定理5.1　$L(V,V')$关于映射的加法与数乘成K上线性空间。

【证明】　为此，仅需验证线性空间定义中的九条公理都满足。

由定义知$\sigma+\tau$，$k\sigma$都是V映入V'的单值映射，且对任何l_1，l_2

$\in K, \alpha, \beta \in V$ 都有

$$(\sigma+\tau)(l_1\alpha+l_2\beta)=\sigma(l_1\alpha+l_2\beta)+\tau(l_1\alpha+l_2\beta)$$
$$=l_1(\sigma+\tau)(\alpha)+l_2(\sigma+\tau)(\beta),$$
$$(k\sigma)(l_1\alpha+l_2\beta)=k(\sigma(l_1\alpha+l_2\beta))$$
$$=l_1(k\sigma)(\alpha)+l_2(k\sigma)(\beta),$$

所以 $\sigma+\tau, k\sigma \in L(V,V')$，即公理$(A_0)$，$(M_0)$满足。

易见，零映射 $0^*: \alpha \to \theta', \forall \alpha \in V$. 就是公理$(A_2)$中所说的零元素，它当然是线性映射。对任一 $\sigma \in L(V,V')$，定义

$$-\sigma:\ \alpha \to -(\sigma(\alpha)), \forall \alpha \in V,$$

它也是线性映射，且 $\sigma+(-\sigma)=(-\sigma)+\sigma=0^*$，所以公理$(A_3)$也满足。

至于公理(A_1)，(M_1)，(M_2)，(M_3)分别归之于 V' 中向量的加法结合律，分配律与数乘结合律，自然都满足。最后，对数 1 显然有

$$(1\cdot\sigma)(\alpha)=1\cdot\sigma(\alpha)=\sigma(\alpha), \forall \alpha \in V,$$

所以公理(M_4)也满足。　　　　证毕。

称 $L(V,V')$为**线性映射空间**。特别，$L(V,V)$称为**线性变换空间**．由 V 中线性函数全体所成的 $L(V, K)$称为 V 的**共轭空间**或**对偶空间**。

对于有限维空间，可得到进一步的结论，即

定理 5.2　设 V 与 V' 分别是数域 K 上的 n 维与 m 维线性空间，$M_{m,n}(K)$ 是 K 上 $m\times n$ 阵关于矩阵的加法与数乘所成的 $m\cdot n$ 维线性空间，则

$$L(V,V')\cong M_{m,n}(K).$$

因而 $L(V, V')$也是 $m\cdot n$ 维线性空间。

【证明】　在 V 与 V' 中取定基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 $\alpha'_1,\alpha'_2,\cdots,\alpha'_m$. 对任一 $\sigma \in L(V,V')$，必可唯一确定一个 $A \in M_{m,n}(K)$与之对应，这个 A 就是 σ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与基 $\alpha'_1, \alpha'_2, \cdots, \alpha'_m$ 下的矩阵，即

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n)=(\alpha'_1, \alpha'_2, \cdots, \alpha'_m)A.$$

于是这就确定了 $L(V, V')$映入 $M_{m,n}(K)$的单值映射

$$\eta:\ \sigma \to A.$$

且由定理 2.4 知 η 是双射。

进一步可证 η 是线性空间 $L(V, V')$ 与 $M_{m,n}(K)$ 之间的同构映射。设对 $\tau \in L(V, V')$ 有

$$\tau(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha'_1, \alpha'_2, \cdots, \alpha'_m)B,$$

则

$$\begin{aligned}
&(\sigma+\tau)(\alpha_1, \alpha_2, \cdots, \alpha_n) \\
&= ((\sigma+\tau)(\alpha_1), (\sigma+\tau)(\alpha_2), \cdots, (\sigma+\tau)(\alpha_n)) \\
&= (\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)) + (\tau(\alpha_1), \tau(\alpha_2), \cdots, \tau(\alpha_n)) \\
&= (\alpha'_1, \alpha'_2, \cdots, \alpha'_m)A + (\alpha'_1, \alpha'_2, \cdots, \alpha'_m)B \\
&= (\alpha'_1, \alpha'_2, \cdots, \alpha'_m)(A+B).
\end{aligned}$$

$$\begin{aligned}
&(k\sigma)(\alpha_1, \alpha_2, \cdots, \alpha_n) \\
&= ((k\sigma)(\alpha_1), (k\sigma)(\alpha_2), \cdots, (k\sigma)(\alpha_n)) \\
&= k(\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)) \\
&= (\alpha'_1, \alpha'_2, \cdots, \alpha'_m)(kA).
\end{aligned}$$

所以

$$\eta(\sigma+\tau) = A + B = \eta(\sigma) + \eta(\tau),$$

$$\eta(k\sigma) = kA = k\eta(\sigma),$$

即 η 是 $L(V, V')$ 到 $M_{m,n}(K)$ 上的同构。 **证毕。**

推论 $L(V, V) \cong M_n(K)$ 为 n^2 维线性空间，

$L(V, K) \cong K_n$ 为 n 维线性空间。

二、线性变换的乘积

关于一般映射的乘积在第九章 § 6 中已有定义。现在我们特别考虑 V 中线性变换的乘积。设 $\sigma, \tau \in L(V, V)$，则对 $\alpha \in V$ 有 $(\sigma\tau)(\alpha) = \sigma(\tau(\alpha))$。

定理 5.3 设 $\sigma, \tau \in L(V, V)$，则 $\sigma\tau \in L(V, V)$。如果 V 是 n 维线性空间，σ 与 τ 在 V 的某个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵分别是 A 与 B，则 $\sigma\tau$ 在此基下的矩阵就是 AB。

【证明】 显然，$\sigma\tau$ 是单值的。对任意 $k_1, k_2 \in K$，$\alpha, \beta \in V$ 有

$$(\sigma\tau)(k_1\alpha + k_2\beta) = \sigma(\tau(k_1\alpha + k_2\beta))$$

$$= \sigma(k_1\tau(\alpha) + k_2\tau(\beta)) = k_1(\sigma\tau)(\alpha) + k_2(\sigma\tau)(\beta),$$

所以 $\sigma\tau \in L(V, V)$。因为

$$
\begin{aligned}
&(\sigma\tau)(\alpha_1, \alpha_2, \cdots, \alpha_n)\\
&=((\sigma\tau)(\alpha_1), (\sigma\tau)(\alpha_2), \cdots, (\sigma\tau)(\alpha_n))\\
&=\sigma(\tau(\alpha_1), \tau(\alpha_2), \cdots, \tau(\alpha_n))\\
&=\sigma((\alpha_1, \alpha_2, \cdots, \alpha_n)B)\\
&=(\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n))B\\
&=(\alpha_1, \alpha_2, \cdots, \alpha_n)AB,
\end{aligned}
$$

所以 $\sigma\tau$ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵恰为 AB。　　　　证毕。

与方阵 A 的幂 A^k 一样，可引进线性变换 σ 的幂。首先定义

$$\sigma^0 = 1^*(\text{单位变换}),$$

$$\sigma^k = \sigma\cdot\sigma^{k-1},\ k\ \text{为正整数},$$

则由定理 5.3 知，若 σ 在某基下的矩阵为 A，则 σ^k 在此基下的矩阵就是 A^k。

我们知道，并非对任意 n 阶阵 A，都有负整数次幂。对线性变换也是如此。

命题 1　K 上线性空间 V 中可逆线性变换的逆变换也是可逆线性变换。

【证明】　设 σ 是可逆线性变换，它的逆变换 σ^{-1} 满足

$$\sigma\cdot\sigma^{-1}=\sigma^{-1}\cdot\sigma=1^*,$$

所以 σ^{-1} 也是可逆变换。对任何 $k_1, k_2\in K$, $\alpha, \beta\in V$ 有

$$
\begin{aligned}
\sigma^{-1}(k_1\alpha+k_2\beta) &= \sigma^{-1}(k_1\sigma\cdot\sigma^{-1}(\alpha)+k_2\sigma\cdot\sigma^{-1}(\beta))\\
&=\sigma^{-1}\sigma(k_1\sigma^{-1}(\alpha)+k_2\sigma^{-1}(\beta))\\
&=k_1\sigma^{-1}(\alpha)+k_2\sigma^{-1}(\beta),
\end{aligned}
$$

所以 σ^{-1} 也是线性变换。　　　　证毕。

于是，对可逆线性变换 σ 可定义负整数次幂

$$\sigma^{-k}=(\sigma^{-1})^k,\ k\ \text{为任意正整数}。$$

在矩阵论中已经知道：若 A 与 B 同为 n 阶阵，则

$$AB=I_n\ \text{当且仅当}\ BA=I_n。$$

因此，A 是可逆阵当且仅当存在 B 使 $AB=I_n$ 或 $BA=I_n$。现在要问：对线性变换来说，情况又是如何呢？

命题 2　设 σ 与 τ 是 n 维线性空间 V 的线性变换，则

$$\sigma\tau = 1^* \text{ 当且仅当 } \tau\sigma = 1^*.$$

【证明】 取定V的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$。考虑定理 5.2 中所建立的 $L(V, V)$到$M_n(K)$的双射

$$\eta: \sigma \to A.$$

设在 η 之下，$\sigma \to A$, $\tau \to B$，则 $\sigma\tau \to AB$, $\tau\sigma \to BA$。但 $1^* \to I_n$，所以由 $\sigma\tau = 1^*$ 和 η 的单值性知 $AB = I_n$，于是 $BA = I_n$。再由 η 是单射知 $\tau\sigma = 1^*$。

同理可知当 $\tau\sigma = 1^*$ 时必有 $\sigma\tau = 1^*$。 证毕。

注意：命题 2 对无限维线性空间未必成立。

例 1 在 $R[x]$中取线性变换σ 和 τ 如下.对 $R[x]$中的多项式

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0$$

定义

$$\sigma(f(x)) = a_nx^{n-1} + a_{n-1}x^{n-2} + \cdots + a_2x + a_1,$$

$$\tau(f(x)) = a_nx^{n+1} + a_{n-1}x^n + \cdots + a_2x^3 + a_1x^2 + a_0x,$$

则易证 σ 与 τ 都是 $R[x]$的线性变换，且 $\sigma\tau = 1^*$，但 $\tau\sigma \neq 1^*$（这里规定 $\sigma\tau$ 是先作用 τ 后作用 σ）。

这是说明有限维与无限维线性空间的区别的又一例子。

习　题

1．设W是线性空间V的线性变换 σ 与 τ 的不变子空间，求证：W也是 $\sigma+\tau$ 与 $\sigma\tau$ 的不变子空间。

2．在R上的多项式空间 $R_{n-1}[x]$中考虑如下三个线性变换：

$$D: \ p(x) \to p'(x),$$

$$\sigma: \ p(x) \to p(x+1) - p(x),$$

$$\tau: \ p(x) \to p(x+1).$$

若 D, σ, τ 在 $R_{n-1}[x]$的基

$$1,\ x,\ \frac{1}{2!}x^2,\ \cdots,\ \frac{1}{(n-1)!}x^{n-1}$$

下的矩阵分别是 N, A, B。求证：$N^n = A^n = 0$，且

$$A = N + \frac{1}{2!}N^2 + \cdots + \frac{1}{(n-1)!}N^{n-1},$$

$$B = I_n + N + \frac{1}{2!}N^2 + \cdots + \frac{1}{(n-1)!}N^{n-1}.$$

3. 称线性空间V的线性变换σ与τ可交换，如果$\sigma\tau = \tau\sigma$。

(i) 取定两个复数z_1和z_2。证明复数域C（作为实线性空间）中如下三个线性变换$\sigma_1, \sigma_2, \sigma_3$两两可交换：

$$\sigma_1: z \to z_1 z,\ \sigma_2: z \to z_2 z,\ \sigma_3: z \to z_1 z_2 z;$$

(ii) 直接求出$\sigma_1, \sigma_2, \sigma_3$在基$1, \sqrt{-1}$下的矩阵$A_1, A_2, A_3$，并验证$A_1A_2 = A_2A_1 = A_3$；

(iii) 应用线性变换与矩阵之间的关系说明$A_1A_2 = A_2A_1 = A_3$。

4．在实多项式空间$R[x]$中取线性变换

$$\sigma(p(x)) = p'(x),\ \tau(p(x)) = xp(x).$$

求证：(i) $\sigma\tau - \tau\sigma = 1^*$；

(ii) $\sigma^n\tau - \tau\sigma^n = n\sigma^{n-1}$, $n = 1, 2, \cdots$。

5．设σ是数域K上线性空间V的线性变换，α是V中某个非零向量使得存在某个正整数m使

$$\sigma^{m-1}(\alpha) \neq \theta \text{ 而 } \sigma^m(\alpha) = \theta.$$

求证：(i) $\alpha, \sigma(\alpha), \cdots, \sigma^{m-1}(\alpha)$在$K$上线性无关；

(ii) 若$m = \mathrm{d}(V)$，则σ在基$\alpha, \sigma(\alpha), \cdots, \sigma^{m-1}(\alpha)$下的矩阵是$n$次幂零阵。

6．设σ是n维线性空间V的线性变换。求证：必存在正整数m使

$$\sigma^m(V) = \sigma^{m+1}(V) = \cdots\cdots.$$

7．设σ与τ是n维线性空间V的线性变换，且$\sigma\tau = 1^*$。证明：σ必是满射，τ必是单射，并由此证明$\tau\sigma = 1^*$。

选 做 题

1．记

$$Z[a, b] = \{f(x) + \sqrt{-1}g(x) \mid f(x), g(x) \in R[a, b]\},$$

这里$R[a, b]$是定义在$[a, b]$上的实连续函数空间。在$Z[a, b]$中定义加法和数量乘法如下：

$$(f(x) + \sqrt{-1}g(x)) + (h(x) + \sqrt{-1}p(x)) = (f(x) + h(x)) + \sqrt{-1}(g(x) + p(x)),$$

$$(k + l\sqrt{-1})(f(x) + \sqrt{-1}g(x)) = (kf(x) - lg(x)) + \sqrt{-1}(lf(x) + kg(x)),$$

求证：(i) $Z[a, b]$是复线性空间；

(ii) $\sigma: f(x)+\sqrt{-1}g(x) \to \int_a^x f(t)\mathrm{d}t+\sqrt{-1}\int_a^x g(t)\mathrm{d}t$ 是 $Z[a, b]$ 中的线性变换；

(iii) 不存在关于 σ 的一维不变子空间。

这说明，定理4.2的推论对这个无限维复线性空间并不成立。

2　设 n 维复线性空间 V 的线性变换 σ 与 τ 可交换，求证：

(i) 若 λ 是 σ 的特征值，则 V_λ 是 τ 的不变子空间；

(ii) σ 与 τ 至少有一个公共的特征向量；

(iii) 如果 n 阶阵 A 与 B 可交换，则必存在非异阵 P 使

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & * \\ & \lambda_2 & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

和

$$P^{-1}BP=\begin{pmatrix}\mu_1 & & * \\ & \mu_2 & \\ & & \ddots & \\ & & & \mu_n\end{pmatrix}$$

同为上三角阵，因而 λ_i，$\mu_i (i=1, 2, \cdots, n)$ 分别为 A 与 B 的特征值；

(iv) $\sigma+\tau$ 的特征值是 $\lambda_i+\mu_i$，$i=1, 2, \cdots, n$。

$\sigma\tau$ 的特征值是 $\lambda_i\mu_i$，$i=1, 2, \cdots, n$。

3．设 $\sigma_1, \sigma_2, \cdots, \sigma_s$ 是线性空间 V 的 s 个两两不同的线性变换，求证：必存在 $\alpha\in V$ 使 $\sigma_1(\alpha), \sigma_2(\alpha), \cdots, \sigma_s(\alpha)$ 也两两不同。

(提示：应用第九章选做题第6题。)

4．设 σ 与 τ 都是线性空间 V 中的幂等线性变换：$\sigma^2=\sigma$，$\tau^2=\tau$，求证：

(i) σ 与 τ 有相同的象空间 $\sigma(V)=\tau(V)$ 的充分必要条件是 $\sigma\tau=\tau$，$\tau\sigma=\sigma$；

(ii) σ 与 τ 有相同的核空间 $\sigma^{-1}(\theta)=\tau^{-1}(\theta)$ 的充分必要条件是 $\sigma\tau=\sigma$，$\tau\sigma=\tau$。

5．设 V 是数域 K 上的 n 维线性空间。如果 V 的线性变换 σ 与 $L(V, V)$ 中的所有可逆线性变换可交换，求证：$\sigma=k1^*$，$k\in K$。

(提示：取具有 n 个线性无关特征向量的某个特殊的可逆线性变换，使 σ 在 V 的基下的矩阵只能是对角阵，然后再取一个可逆线性变换，使 σ 在 V 的基下的矩阵只能是数量阵。)

6. 设V是数域K上n维线性空间，记$V^*=L(V, K)$，求证：

(i) $d(V)=d(V^*)$；

(ii) 对于V的每一个基$\alpha_1, \alpha_2, \cdots, \alpha_n$，在$V^*$中必存在唯一的基$f_1, f_2, \cdots, f_n$使

$$f_i(\alpha_j)=\delta_{ij}, \quad i, j=1, 2, \cdots, n;$$

(iii) 记$V^{**}=(V^*)^*$，即V^{**}是$L(V, K)$上的线性函数全体所成的线性空间，则$V^{**}\cong V$（称为对偶原理）；

(iv) 对V^*中任一基$f_1, f_2, \cdots, f_n$，必存在V的唯一的基$\alpha_1, \alpha_2, \cdots, \alpha_n$使

$$\bar{\alpha}_i(f_j)=\delta_{ij}, \ i, \quad j=1, 2, \cdots, n,$$

这里$\bar{\alpha}_i$是$\alpha_i \in V$在V^{**}中的同构象。

（提示：对(ii)，只要用定理2.2便可证得。对(iii)，任取$\alpha \in V$，作单值映射：$\bar{\alpha}(f)=f(\alpha), \forall f \in V^*$，即可证得所需结论。而由(ii)与(iii)即得(iv)。）

第十一章 欧 氏 空 间

本章介绍欧氏空间的一些基本概念以及欧氏空间中两个重要的线性变换——自共轭变换与正交变换，并介绍了欧氏空间及其线性变换的应用。

§1 内积、Gram 矩阵的半正定性

一、内积、欧氏空间

线性空间是通常三维几何(实)空间的抽象，在三维几何空间中，除了向量的加法、实数与向量的数乘运算外，还有两个向量**内积**的概念，由这个度量概念可以获得不少有用的几何性质，对一般的实线性空间，我们自然也希望引进类似的度量概念，以便进一步得到空间更多的性质．于是有下面的

定义 设 V 是实数域 R 上的线性空间，对 V 中任意两个向量 α 与 β，定义唯一的一个实数，记为 (α, β)，如果 (α, β) 满足：

(i) $(\alpha, \beta)=(\beta, \alpha)$， $\forall \alpha, \beta \in V$．

(ii) $(\alpha+\beta, \gamma)=(\alpha, \gamma)+(\beta, \gamma)$，$\forall \alpha, \beta, \gamma \in V$．

(iii) $(k\alpha, \beta)=k(\alpha, \beta)$， $\forall \alpha, \beta \in V; k \in K$．

(iv) $(\alpha, \alpha) \geqslant 0$，而等号成立的充要条件是：$\alpha=\theta$，则称 (α, β) 是 α 与 β 的**内积**．

称带有内积的 R 上的线性空间 V 为 Euclid **空间**．以后简称 V 为**欧氏空间**．

例 1 对 n 维列向量空间 $R^{(n)}$ 中任意两个向量 $\alpha=(a_1, a_2, \cdots, a_n)'$ 与 $\beta=(b_1, b_2, \cdots, b_n)'$，定义

$$(\alpha, \beta) = \alpha'\beta = \sum_{i=1}^{n} a_i b_i,$$

则易证(α, β)满足条件(i)至(iv)，所以(α, β)是一内积，于是带有内积 $\alpha'\beta$ 的 $R^{(n)}$ 是一个有限维(n 维)欧氏空间，当 $n=3$，这就是通常的三维几何空间。

例 2 对 $R[a, b]$中任意两个实变实值函数 f、g，定义

$$(f, g) = \int_a^b f(x)g(x)\mathrm{d}x,$$

则易证(f, g)是一内积，所以带有内积 (f, g) 的 $R[a, b]$是一个无限维欧氏空间。

内积有下列简单性质

(i) $(\theta, \beta) = 0$, $\qquad \forall \beta \in V$,

(ii) $(\alpha, \beta+\gamma) = (\alpha, \beta) + (\alpha, \gamma)$, $\forall \alpha, \beta, \gamma \in V$,

一般成立：

$$\left(\alpha, \sum_{i=1}^{s} \beta_i\right) = \sum_{i=1}^{s} (\alpha, \beta_i), \ \forall \beta_i \in V, \tag{1}$$

(iii) $(\alpha, k\beta) = k(\alpha, \beta)$, $\forall \alpha, \beta \in V; k \in K$,

这三个性质都是易证的：(i) 是由于

$$(\alpha, \beta) = (\alpha+\theta, \beta) = (\alpha, \beta) + (\theta, \beta)$$

而得到，(ii)是由于

$$(\alpha, \beta+\gamma) = (\beta+\gamma, \alpha) = (\beta, \alpha) + (\gamma, \alpha) = (\alpha, \beta) + (\alpha, \gamma),$$

一般式子(1)只要用(ii)及归纳法便可证得，(iii)是由于

$$(\alpha, k\beta) = (k\beta, \alpha) = k(\beta, \alpha) = k(\alpha, \beta).$$

二、Cauchy-Буняковский 不等式

欧氏空间中的一个基本结果是下面的不等式：

定理 1.1 对欧氏空间 V 中任意两个向量 α, β，恒有

$$(\alpha, \beta)^2 \leqslant (\alpha, \alpha)(\beta, \beta) \tag{2}$$

而等号成立的充要条件是，α 与 β 线性相关。

【证明】 对任意实变数 x_1、x_2，由于内积的条件(iv)可知

$$(x_1\alpha + x_2\beta, x_1\alpha + x_2\beta) \geqslant 0. \tag{3}$$

应用内积的条件及性质，上式可以化为

$$(\alpha,\alpha)x_1^2+2(\alpha,\beta)x_1x_2+(\beta,\beta)x_2^2\geqslant 0.$$

这说明上面的二次型是半正定的，故

$$\begin{vmatrix}(\alpha,\alpha) & (\alpha,\beta)\\ (\beta,\alpha) & (\beta,\beta)\end{vmatrix}\geqslant 0 \quad (\because(\alpha,\beta)=(\beta,\alpha)). \tag{4}$$

这证明了(2)式。

如果α与β线性相关：$\beta=k\alpha$，则(2)的等号显然成立。反之，如果(2)的等号成立，而α与β线性无关，则对非零数x_1、x_2，必有$x_1\alpha+x_2\beta\neq\theta$。于是二次型(3)是正定二次型，从而(4)式中等号应去掉，也即$(\alpha,\alpha)(\beta,\beta)>(\alpha,\beta)^2$，此与假设（等号成立）相矛盾，故$\alpha$与$\beta$线性相关。　　　　证毕。

称(2)式为 Cauchy–Буняковский 不等式，由这个不等式显然可以推得下面两个不等式。

推论 1　设$A=(a_{ij})_{n\times n}$是正定阵，$x=(x_1,x_2,\cdots,x_n)'$，$y=(y_1,y_2,\cdots,y_n)'$是实向量，则

$$\left(\sum_{i,j=1}^{n}a_{ij}x_iy_j\right)^2\leqslant\left(\sum_{i,j=1}^{n}a_{ij}x_ix_j\right)\left(\sum_{i,j=1}^{n}a_{ij}y_iy_j\right) \tag{5}$$

或者

$$(x'Ay)^2\leqslant(x'Ax)(y'Ay). \tag{6}$$

特别，当$A=I_n$，则又得到通常的 Cauchy–Schwarz 不等式。

推论 2　设$f(x)$，$g(x)$是$R[a,b]$中任意两个元素，则

$$\left(\int_a^b f(x)g(x)\mathrm{d}x\right)^2\leqslant\left(\int_a^b f^2(x)\mathrm{d}x\right)\left(\int_a^b g^2(x)\mathrm{d}x\right). \tag{7}$$

三、向量的长度，两个向量的交角

定义　设α是欧氏空间中任一向量，称

$$\|\alpha\|=(\alpha,\alpha)^{1/2}$$

为α的**长度**或**范数**。

由内积的定义显然可得：

(i) $\|\alpha\|\geqslant 0$，而等号成立的充要条件是，$\alpha=\theta$；

(ii) $\|k\alpha\|=k\|\alpha\|$，$\forall k\in K$；

(iii) $\|\alpha+\beta\|\leqslant\|\alpha\|+\|\beta\|$, $\forall\alpha,\beta\in V$（称为**“三角形不等式”**）。

(iii)是由于应用内积定义以及 Cauchy–Буняковский 不等式可得，即

$$\begin{aligned}\|\alpha+\beta\|^2&=(\alpha+\beta,\alpha+\beta)=(\alpha,\alpha)+2(\alpha,\beta)+(\beta,\beta)\\&\leqslant(\alpha,\alpha)+2(\alpha,\alpha)^{1/2}(\beta,\beta)^{1/2}+(\beta,\beta)\\&=(\|\alpha\|+\|\beta\|)^2,\end{aligned}$$

所以(iii)是正确的。

由向量的长度可以定义两个向量的“距离”，对欧氏空间 V 中任意两个向量 α、β，定义

$$\rho(\alpha,\beta)=\|\alpha-\beta\|$$

称 $\rho(\alpha,\beta)$ 为 α 与 β 的**距离**，由向量长度的定义及三角形不等式，易知成立

(i) $\rho(\alpha,\alpha)=0$；

(ii) $\rho(\alpha,\beta)=\rho(\beta,\alpha)$；

(iii) $\rho(\alpha,\beta)\leqslant\rho(\alpha,\gamma)+\rho(\gamma,\beta)$

（“三角形”两“边”之和大于第三“边”）。

下面引进两个向量的交角。对欧氏空间 V 中任意两个非零向量 α、β，定义

$$\langle\alpha,\beta\rangle=\arccos\frac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|}\tag{8}$$

为 α 与 β 的**交角**，由 Cauchy–Буняковский 不等式，$-1\leqslant\frac{(\alpha,\beta)}{\|\alpha\|\cdot\|\beta\|}\leqslant1$，故 $|\cos\langle\alpha,\beta\rangle|\leqslant1$，这说明定义 (8) 是合理的．且容易看出，(8)式是通常三维几何空间中两个向量夹角的概念的推广。

定义 如果 $\langle\alpha,\beta\rangle=\frac{\pi}{2}$，则称 α 与 β 是**正交的**，以 $\alpha\perp\beta$ 记之。

命题 1 $\alpha\perp\beta$ 的充要条件是．$(\alpha,\beta)=0$。

命题 2（**商高定理**或**勾股定理**）欧氏空间中两个向量 α 与 β 正交的充要条件是

$$\|\alpha\|^2+\|\beta\|^2=\|\alpha+\beta\|^2.\tag{9}$$

命题 1 是显然的，命题 2 由等式：

$$\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2+2(\alpha, \beta)$$

即可看出。

四、对称双线性函数，度量矩阵

由内积的定义的第(iii)、(ii)两条可知，对任一暂时固定的 β，(α, β)是关于 α 的线性(实)函数，再由性质(ii)、(iii)，对任一暂时固定的 α，(α, β)是关于 β 的线性函数，于是我们称(α, β)为**双线性函数**。再由内积定义(i)的对称性质，所以进一步称(α, β)为**对称双线性(泛)函数**。

对 n 维欧氏空间来说，取定 V 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$，则

$$\alpha=\sum_{i=1}^{n} x_i\alpha_i, \quad \beta=\sum_{j=1}^{n} y_j\alpha_j,$$

于是对称双线性函数(α, β)可写成

$$\begin{aligned}(\alpha, \beta)&=\left(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{n} y_j\alpha_j\right)=\sum_{i,j=1}^{n}(\alpha_i, \alpha_j)x_iy_j\\&=\sum_{i,j=1}^{n} a_{ij}x_iy_j=x'Ay,\end{aligned}$$

其中记：

$$a_{ij}=(\alpha_i, \alpha_j), \quad A=(a_{ij})_{n\times n},$$

$$x=(x_1, x_2, \cdots, x_n)', \quad y=(y_1, y_2, \cdots, y_n)',$$

称 A 为基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的**度量矩阵**。由于 $(\alpha_i, \alpha_j)=(\alpha_j, \alpha_i)$，故 A 是实对称阵，$x'Ay$ 是对称双线性型。

命题 3 度量矩阵是正定阵。

这是因为，$(\alpha, \alpha)=\sum_{i,j=1}^{n} a_{ij}x_ix_j\geqslant 0$，而等号成立的充要条件是 $\alpha=\theta$，即 $x=0$，所以 A 是正定阵。

对实数域上的 n 维线性空间 V，在取定的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 以及给定的 n 阶正定阵 $A=(a_{ij})_{n\times n}$ 后，对 V 中任意向量 α, β，就可定义

$$(\alpha, \beta)=\sum_{i,j=1}^{n} a_{ij}x_iy_j,$$

其中 $\alpha=\sum_{i=1}^{n} x_i\alpha_i$，$\beta=\sum_{j=1}^{n} y_j\alpha_j$。则易证$(\alpha, \beta)$是一内积，从而$V$成为$n$维欧氏空间，而$A$就是这个基的度量矩阵。故得

命题 4 设V是实数域上的n维线性空间，则对任一n阶正定阵A，必可定义V的一个内积，使V成为n维欧氏空间。

命题 5 n维欧氏空间V的不同基的度量矩阵是合同的。

【证明】 对V的不同的基$\alpha_1, \alpha_2, \cdots, \alpha_n$与$\beta_1, \beta_2, \cdots, \beta_n$及任意实向量$W=(w_1, w_2, \cdots, w_n)'$，$z=(z_1, z_2, \cdots, z_n)'$，作向量：

$$\alpha=(\beta_1, \beta_2, \cdots, \beta_n)\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}, \beta=(\beta_1, \beta_2, \cdots, \beta_n)\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}.$$

设α、β在基$\alpha_1, \alpha_2, \cdots, \alpha_n$下的表示式是：

$$\alpha=(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \beta=(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

又设$\alpha_1, \alpha_2, \cdots, \alpha_n$到$\beta_1, \beta_2, \cdots, \beta_n$的过渡矩阵是$P$，即

$$(\beta_1, \beta_2, \cdots, \beta_n)=(\alpha_1, \alpha_2, \cdots, \alpha_n)P.$$

于是

$$(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\alpha=(\alpha_1, \alpha_2, \cdots, \alpha_n)P\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix},$$

$$(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=\beta=(\alpha_1, \alpha_2, \cdots, \alpha_n)P\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}.$$

由于$\alpha_1, \alpha_2, \cdots, \alpha_n$线性无关，故由上两式即得

$$x=PW, \quad y=PZ, \tag{10}$$

其中 $x=(x_1, x_2, \cdots, x_n)'$，$y=(y_1, y_2, \cdots, y_n)'$。

设基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\beta_1, \beta_2, \cdots, \beta_n$ 的度量矩阵分别是 A 与 B. 则

$$(\alpha, \beta) = x'Ay = W'BZ.$$

由(10)式,上式可写成

$$W'(P'AP)Z = W'BZ. \tag{11}$$

因为 W, Z 可以取任意实向量,故由(11)式可得

$$B = P'AP.$$

证毕.

五、Gram 矩阵,广义 Hadamard 不等式

度量矩阵是下面被称为 Gram 矩阵的一个特例.

定义 设 V 是任意一个欧氏空间, $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 V 中任意 s 个向量,称 s 阶阵

$$G(\alpha_1, \alpha_2, \cdots, \alpha_s) = \begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) \cdots (\alpha_1, \alpha_s) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) \cdots (\alpha_2, \alpha_s) \\ \cdots\cdots & \cdots\cdots\cdots\cdots \\ (\alpha_s, \alpha_1) & (\alpha_s, \alpha_2) \cdots (\alpha_s, \alpha_s) \end{pmatrix}$$

为($\alpha_1, \alpha_2, \cdots, \alpha_s$ 的)**Gram 矩阵**,而称 $|G(\alpha_1, \alpha_2, \cdots, \alpha_s)|$ 为 **Gram 行列式.**

例 3 $R^{(n)}$ 的 s 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 对内积 $\alpha'\beta$ 来说的Gram 矩阵是

$$G(\alpha_1, \alpha_2, \cdots, \alpha_s) = \begin{pmatrix} \alpha_1'\alpha_1 & \alpha_1'\alpha_2 \cdots \alpha_1'\alpha_s \\ \alpha_2'\alpha_1 & \alpha_2'\alpha_2 \cdots \alpha_2'\alpha_s \\ \cdots\cdots & \cdots\cdots\cdots \\ \alpha_s'\alpha_1 & \alpha_s'\alpha_2 \cdots \alpha_s'\alpha_s \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_1' \\ \alpha_2' \\ \vdots \\ \alpha_s' \end{pmatrix} (\alpha_1, \alpha_2, \cdots, \alpha_s) = B'B,$$

其中 $B = (\alpha_1, \alpha_2, \cdots, \alpha_s)$是 $n \times s$ 阵.

例 4 $R[a, b]$中任意 s 个实变实值函数 $f_1, f_2, \cdots, f_s$ 对内积 $\int_a^b f(x)g(x)dx$ 来说的 Gram 矩阵是

$$G(f_1, f_2, \cdots, f_s)$$

$$=\begin{pmatrix} \int_a^b f_1^2(x)\mathrm{d}x & \int_a^b f_1(x)f_2(x)\mathrm{d}x\cdots & \int_a^b f_1(x)f_s(x)\mathrm{d}x \\ \int_a^b f_2(x)f_1(x)\mathrm{d}x & \int_a^b f_2^2(x)\mathrm{d}x\cdots\cdots & \int_a^b f_2(x)f_s(x)\mathrm{d}x \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \int_a^b f_s(x)f_1(x)\mathrm{d}x & \int_a^b f_s(x)f_2(x)\mathrm{d}x\cdots & \int_a^b f_s^2(x)\mathrm{d}x \end{pmatrix}$$

Gram 矩阵有下面十分重要的性质，它是定理 1.1 的推广。

定理 1.2 设 V 是任意一个欧氏空间，则 V 中任意 s 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的 Gram 矩阵 $G(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 必是半正定阵，而 $G(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是正定阵的充要条件是，$\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关。

【证明】 因为

$$\sum_{i,j=1}^{s} (\alpha_i, \alpha_j)x_i x_j = \left(\sum_{i=1}^{s} x_i\alpha_i, \sum_{i=1}^{s} x_i\alpha_i\right) \geqslant 0, \tag{12}$$

故上述二次型是半正定二次型，也即 $G(\alpha_1, \alpha_2, \cdots, \alpha_s) = ((\alpha_i, \alpha_j))_{s\times s}$ 是半正定阵。

由于(12)式等号成立的充要条件是，$\sum_{i=1}^{s} x_i\alpha_i = \theta$，故当 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关时，对任一非零实向量 $(x_1, x_2, \cdots, x_s)'$，恒有 $\sum_{i=1}^{s} x_i\alpha_i \neq \theta$，故 $\sum_{i,j=1}^{s} (\alpha_i, \alpha_j)x_i x_j > 0$，即 $G(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是正定阵。反之，若 $G(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 是正定阵，则对任何非零实向量 $(x_1, x_2, \cdots, x_s)'$ 恒有 $\sum_{i=1}^{s} (\alpha_i, \alpha_j)x_i x_j > 0$，如果 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关，则存在不全为零的数 $x_1, x_2, \cdots, x_n$ 使 $\sum_{i=1}^{s} x_i\alpha_i = \theta$，此时 $\sum_{i=1}^{s}(\alpha_i, \alpha_j)x_i x_j = 0$，此为矛盾，故 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关。证毕。

由定理 1.2 显然可得以下两个推论

推论 1 欧氏空间 V 中任意 s 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的 Gram 矩阵的任何主子式都是非负数，当且仅当 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关时，$|G(\alpha_1, \alpha_2, \cdots, \alpha_s)| > 0$。

推论 2 （广义 Hadamard 不等式）对欧氏空间中任意 s 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$，必有

$$|G(\alpha_1, \alpha_2, \cdots, \alpha_s)| \leqslant \prod_{i=1}^{s}(\alpha_i, \alpha_i), \tag{13}$$

而等号成立的充要条件是$(\alpha_i, \alpha_j)=0$, $i \neq j$, $i, j=1, 2, \cdots, s$（即 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 两两正交）。

由于 $G(\alpha_1, \alpha_2, \cdots, \alpha_s)$ 的半正定性，使我们可以充分运用半正定阵以及正定阵的整套理论去获得任意欧氏空间中一些向量的有用结果。例如，由半正定阵（正定阵）乘积的 Schur 定理显然可得

推论 3 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是任一欧氏空间中的任意 s 个向量，则对任何正整数 k，方阵

$$A(k)=\begin{pmatrix}(\alpha_1, \alpha_1)^k & (\alpha_1, \alpha_2)^k \cdots (\alpha_1, \alpha_s)^k \\ (\alpha_2, \alpha_1)^k & (\alpha_2, \alpha_2)^k \cdots (\alpha_2, \alpha_s)^k \\ \cdots\cdots\cdots\cdots & \\ (\alpha_s, \alpha_1)^k & (\alpha_s, \alpha_2)^k \cdots (\alpha_s, \alpha_s)^k\end{pmatrix}$$

必是半正定阵。当 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关时，则 $A(k)$ 必是正定阵。

例 5 设 $f_1(x), \cdots, f_s(x)$ 是$[a, b]$上 s 个实变实值函数，则

$$\begin{vmatrix}\int_a^b f_1^2(x)\mathrm{d}x & \int_a^b f_1(x)f_2(x)\mathrm{d}x \cdots & \int_a^b f_1(x)f_s(x)\mathrm{d}x \\ \int_a^b f_2(x)f_1(x)\mathrm{d}x & \int_a^b f_2^2(x)\mathrm{d}x \cdots & \int_a^b f_2(x)f_s(x)\mathrm{d}x \\ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ \int_a^b f_s(x)f_1(x)\mathrm{d}x & \int_a^b f_s(x)f_2(x)\mathrm{d}x \cdots & \int_a^b f_s^2(x)\mathrm{d}x\end{vmatrix}$$

$$\leqslant \prod_{i=1}^{s}\left(\int_a^b f_i^2(x)\mathrm{d}x\right).$$

例 6 设 $A=(\alpha_1, \alpha_2, \cdots, \alpha_n)$是 n 阶实方阵，则

$$|A|^2=|A'A|=|G(\alpha_1, \alpha_2, \cdots, \alpha_n)| \leqslant \prod_{i=1}^{n}\alpha_i'\alpha_i=\prod_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2.$$

这就是通常的 Hadamard 不等式。

例 7 设 $f_1(x), f_2(x), f_3(x)$是实变实值函数，则

$$\begin{vmatrix} \left(\int_a^b f_1^2(x)\mathrm{d}x\right)^3 & \left(\int_a^b f_1(x)f_2(x)\mathrm{d}x\right)^3 & \left(\int_a^b f_1(x)f_3(x)\mathrm{d}x\right)^3 \\ \left(\int_a^b f_2(x)f_1(x)\mathrm{d}x\right)^3 & \left(\int_a^b f_2^2(x)\mathrm{d}x\right)^3 & \left(\int_a^b f_2(x)f_3(x)\mathrm{d}x\right)^3 \\ \left(\int_a^b f_3(x)f_1(x)\mathrm{d}x\right)^3 & \left(\int_a^b f_3(x)f_2(x)\mathrm{d}x\right)^3 & \left(\int_a^b f_3^2(x)\mathrm{d}x\right)^3 \end{vmatrix}$$

$\geqslant 0$。

习　题

1．分别举出不同于课文上的有限维与无限维欧氏空间的例子。

2．设A是n阶实对称阵，α与β是$R^{(n)}$中任两向量，求证：

(i) 当A是半正定阵时，则$(\alpha'A\beta)^2\leqslant(\alpha'A\alpha)(\beta'A\beta)$；

(ii) 当A是正定阵时，则$(\alpha'\beta)^2\leqslant(\alpha'A\alpha)(\beta'A^{-1}\beta)$。

(提示：由$A=Q[\lambda_1, \cdots, \lambda_n]Q'$及$A$的正定条件，把$\alpha'\beta$写成：$\alpha'\beta=((Q[\lambda_1^{1/2}, \cdots, \lambda_n^{1/2}]Q')\alpha)'((Q[\lambda_1^{-1/2}, \cdots, \lambda_n^{-1/2}]Q')(Q\beta))$，(ii)即得证。)

3．对欧氏空间中任意两个向量α、β，求证：

(i) $\|\alpha-\beta\|\geqslant\|\alpha\|-\|\beta\|$；

(ii) $\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2\|\alpha\|^2+2\|\beta\|^2$(平行四边形法则)；

(iii) $(\alpha, \beta)=\dfrac{1}{4}(\|\alpha+\beta\|^2-\|\alpha-\beta\|^2)$；

(iv) $\alpha\perp\beta\Longleftrightarrow\|\alpha+\beta\|=\|\alpha-\beta\|$。

4．设α、β、γ是通常三维(欧氏)几何空间的三个线性无关向量，求证：

(i) $|G(\alpha, \beta, \gamma)|$是以$\alpha$、$\beta$、$\gamma$为"棱"的平行六面体的体积的平方；

(ii) 以α、β、γ为棱的平行六面体的体积不大于各向量棱长的乘积，而等号成立的充要条件是，六面体是长方体。

5．设α、β、γ是任一欧氏空间中的线性无关向量，求证：

$$\begin{pmatrix} (\alpha, \alpha)^2 & (\alpha, \beta)^2 & (\alpha, \gamma)^2 \\ (\beta, \alpha)^2 & (\beta, \beta)^2 & (\beta, \gamma)^2 \\ (\gamma, \alpha)^2 & (\gamma, \beta)^2 & (\gamma, \gamma)^2 \end{pmatrix}$$

的任何主子式全大于零。

6．设$f_1(x)$，$f_2(x)$，$f_3(x)$是$R[a, b]$的线性无关向量，求证：

$$\max_{1\leqslant i\leqslant 3}\int_a^b f_i^2(x)dx\geqslant\int_a^b f_j(x)f_k(x)dx,\ j\neq k,\ j,\ k=1,\ 2,\ 3.$$

7．设$\alpha_1, \alpha_2, \cdots, \alpha_s$是欧氏空间中任意$s$个向量，对任何$k$，$1\leqslant k\leqslant s$，

求证：

$$|G(\alpha_1, \alpha_2, \cdots, \alpha_s)| \leqslant |G(\alpha_1, \alpha_2, \cdots, \alpha_k)| \cdot |G(\alpha_{k+1}, \alpha_{k+2}, \cdots, \alpha_s)|,$$

当 $s=3, k=2$ 时，试述上面不等式的几何意义。

8. 证明下列不等式

$$\begin{vmatrix} \int_{-1}^{1} x^{2n-2}(x^2+1)\mathrm{d}x & \int_{-1}^{1} x^{2n-3}(x^2+1)\mathrm{d}x \cdots \int_{-1}^{1} x^{n-1}(x^2+1)\mathrm{d}x \\ \int_{-1}^{1} x^{2n-3}(x^2+1)\mathrm{d}x & \int_{-1}^{1} x^{2n-4}(x^2+1)\mathrm{d}x \cdots \int_{-1}^{1} x^{n-2}(x^2+1)\mathrm{d}x \\ \cdots\cdots & \cdots\cdots \\ \int_{-1}^{1} x^{n-1}(x^2+1)\mathrm{d}x & \int_{-1}^{1} x^{n-2}(x^2+1)\mathrm{d}x \cdots \int_{-1}^{1} x^{0}(x^2+1)\mathrm{d}x \end{vmatrix}$$

$$> \begin{vmatrix} \int_{-1}^{1} x^{2n}\mathrm{d}x & \int_{-1}^{1} x^{2n-1}\mathrm{d}x \cdots \int_{-1}^{1} x^{n+1}\mathrm{d}x \\ \int_{-1}^{1} x^{2n-1}\mathrm{d}x & \int_{-1}^{1} x^{2n-2}\mathrm{d}x \cdots \int_{-1}^{1} x^{n}\mathrm{d}x \\ \cdots\cdots & \cdots\cdots \\ \int_{-1}^{1} x^{n+1}\mathrm{d}x & \int_{-1}^{1} x^{n}\mathrm{d}x \cdots \int_{-1}^{1} x^{2}\mathrm{d}x \end{vmatrix}.$$

§2 正交向量组、欧氏空间的同构

一、正交向量组、正交化方法

定义 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是欧氏空间 V 的一组非零向量，如果 $(\alpha_i, \alpha_j)=0$, $i\neq j$; $i, j=1, 2, \cdots, s$, 则称 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是**正交向量组**。如果正交向量组的每一个向量的长度都是 1，则称它们为**标准正交向量组**，或称为**么正向量组**。

今后常以 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_s$ 表示么正向量组，显然 $(\varepsilon_i, \varepsilon_j)=\delta_{ij}$, $i\neq j$; $i, j=1, 2, \cdots, s$。

例 1 向量组

$$\alpha_1=(1, -2, 0, 0)'$$
$$\alpha_2=(2, 1, 1, 1)'$$
$$\alpha_3=(0, 0, 1, -1)'$$

是 $R^{(4)}$ 的正交向量组（内积的定义是：$(\alpha, \beta)=\alpha'\beta$），而

$$\varepsilon_1=\left(\frac{1}{\sqrt{5}}, \frac{-2}{\sqrt{5}}, 0, 0\right)'$$

$$\varepsilon_2=\left(\frac{-2}{\sqrt{7}},\ \frac{1}{\sqrt{7}},\ \frac{1}{\sqrt{7}},\ \frac{1}{\sqrt{7}}\right)'$$

$$\varepsilon_3=\left(0,\ 0,\ \frac{1}{\sqrt{2}},\ \frac{-1}{\sqrt{2}}\right)'$$

是标准正交组。且 $\varepsilon_i=\frac{\alpha_i}{\|\alpha_i\|}$，即把 α_i "单位化"成 ε_i，$i=1, 2, 3$，即得么正组。

例 2 无限维欧氏空间 $R[0, 2\pi]$的向量组：

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cdots,\ \cos tx,\ \sin tx$$

是正交向量组$\left(\text{内积是：}(f, g)=\int_0^{2\pi}f(x)g(x)\mathrm{d}x\right)$，这是因为：当 $k\neq l$ 时，

$$\int_0^{2\pi}\sin kx\sin lx\,\mathrm{d}x=0,\int_0^{2\pi}\cos kx\cos lx\mathrm{d}x=0,$$

且对任何 k, l，恒有：

$$\int_0^{2\pi}\sin kx\cos lx\mathrm{d}x=0,$$

$k, l=0, 1, 2, \cdots, t$。而向量组

$$\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\cos x,\ \frac{1}{\sqrt{\pi}}\sin x,\ \frac{1}{\sqrt{\pi}}\cos 2x,\ \frac{1}{\sqrt{\pi}}\sin 2x,\ \cdots,$$

$$\frac{1}{\sqrt{\pi}}\cos tx,\ \frac{1}{\sqrt{\pi}}\sin tx$$

是标准正交向量组。

命题 1 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是欧氏空间 V 的正交向量组，则 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 必线性无关。

【证明】 因由 $\sum_{i=1}^{s}k_i\alpha_i=\theta$ 可得

$$k_j(\alpha_j, \alpha_j)=\left(\sum_{i=1}^{s}k_i\alpha_i,\ \alpha_j\right)=(\theta, \alpha_j)=0,\quad 1\leqslant j\leqslant s,$$

由于 $\alpha_j\neq\theta$，故$(\alpha_j, \alpha_j)\neq 0$，故由上式可知，对任何 j，恒有 $k_j=0$，$j=1, 2, \cdots, s$，所以 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关。　　证毕。

定理 2.1 任何欧氏空间 V 必存在正交向量组。

【证明】 下面所用的构造性方法，其思想完全由平面上的两个线性无关向量作出正交向量组以及空间中三个线性无关向量作

出两两正交向量的几何方法抽象出来，具体构造如下：取 V 中某 s 个线性无关的向量 $\beta_1, \beta_2, \cdots, \beta_s$，令

$$\begin{aligned}
\alpha_1 &= \beta_1 \\
\alpha_2 &= \beta_2 - \frac{(\beta_2, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 \\
\alpha_3 &= \beta_3 - \frac{(\beta_3, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 - \frac{(\beta_3, \alpha_2)}{(\alpha_2, \alpha_2)}\alpha_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (1) \\
\alpha_s &= \beta_s - \frac{(\beta_s, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 - \frac{(\beta_s, \alpha_2)}{(\alpha_2, \alpha_2)}\alpha_2 - \cdots - \frac{(\beta_s, \alpha_{s-1})}{(\alpha_{s-1}, \alpha_{s-1})}\alpha_{s-1}
\end{aligned}$$

则易证 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 都是非零向量，因为：由(1)的第一个式子与第二个式子可知 α_2 是 β_1 与 β_2 的线性组合，再由(1)的第一、二、三这三个式子可知 α_3 是 $\beta_1, \beta_2, \beta_3$ 的线性组合，依次推下去可知，对任一 i，α_i 都是 $\beta_1, \beta_2, \cdots, \beta_i$ 的线性组合，且

$$\alpha_i = \beta_i + k_1\beta_1 + k_2\beta_2 + \cdots + k_{i-1}\beta_{i-1}, \quad 1 \leqslant i \leqslant s, \qquad (2)$$

因为 $\beta_1, \beta_2, \cdots, \beta_s$ 线性无关，故由(2)式显然可知 $\alpha_i \neq \theta$，$i = 1, 2, \cdots, s$。

今证 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 两两正交，对 s 用归纳法，由于

$$(\alpha_1, \alpha_2) = \left(\beta_1, \beta_2 - \frac{(\beta_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1\right) = (\beta_1, \beta_2) - (\beta_2, \beta_1) = 0,$$

故结论对 $s = 2$ 正确。设当 $s-1$ 时，$\alpha_1, \alpha_2, \cdots, \alpha_s$ 中任意 $s-1$ 个向量两两正交，要证这 s 个向量中的任意两个两两正交，事实上只要证 α_s 与 $\alpha_1, \alpha_2, \cdots, \alpha_{s-1}$ 中的任一向量 α_j $(1 \leqslant j \leqslant s-1)$ 正交即可。由于

$$\begin{aligned}
(\alpha_s, \alpha_j) &= (\beta_s, \alpha_j) - \sum_{i=1}^{s-1} \frac{(\beta_s, \alpha_i)}{(\alpha_i, \alpha_i)}(\alpha_i, \alpha_j) \\
&= (\beta_s, \alpha_j) - (\beta_s, \alpha_j) = 0,
\end{aligned}$$

故 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是正交向量组。　　　　证毕。

用(1)式把 s 个线性无关向量构造成 s 个向量所成的正交向量组的方法叫作 Gram-Schmidt **正交化方法**。

例 3　把欧氏空间 $R^{(n)}$（内积是 $\alpha'\beta$）中两个向量：$\beta_1 = (1, 1, \cdots, 1)'$，$\beta_2 = (1, -1, 1, \cdots, 1)'$ 构造为正交向量组。

【解】 由(1)式得到：

$$\alpha_1=\beta_1=(1,1,\cdots,1)',$$

$$\alpha_2=\beta_2-\frac{(\beta_2,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1=\beta_2-\frac{(\beta_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$$

$$=\left(\frac{2}{n}-\frac{2(n-1)}{n},\ \frac{2}{n},\cdots,\ \frac{2}{n}\right)$$

推论 1 n 维欧氏空间 V 必存在标准正交向量组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$。

【证明】 任取 V 的一个基 $\beta_1,\beta_2,\cdots,\beta_n$，由定理 2.1 可把这 n 个向量正交化为 $\alpha_1,\alpha_2,\cdots,\alpha_n$，使 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的正交向量组，再把 α_i 单位化，令 $\varepsilon_i=\dfrac{\alpha_i}{\|\alpha_i\|}$，$i=1,2,\cdots,n$，即得 V 的标准正交向量组 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$。 证毕。

由命题 1 可知，推论 1 得出的 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关，故 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的基，称为**标准正交基**，或称为**么正基**。

例 4 求证：n 阶正交阵的 n 个列向量都可作为 $R^{(n)}$ 的标准正交基。反之，$R^{(n)}$ 的标准正交基所成的 n 阶阵必是正交阵。

【证明】 把 A 按列分块：$A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$，因为 A 是正交阵，故

$$A'A=\begin{pmatrix}\alpha_1'\\\alpha_2'\\\vdots\\\alpha_n'\end{pmatrix}(\alpha_1,\alpha_2,\cdots,\alpha_n)$$

$$=\begin{pmatrix}\alpha_1'\alpha_1 & \alpha_1'\alpha_2\cdots\alpha_1'\alpha_n\\\alpha_2'\alpha_1 & \alpha_2'\alpha_2\cdots\alpha_2'\alpha_n\\\cdots\cdots\cdots\cdots & \cdots\\\alpha_n'\alpha_1 & \alpha_n'\alpha_2\cdots\alpha_n'\alpha_n\end{pmatrix}=I_n,$$

所以

$$(\alpha_i,\alpha_j)=\alpha_i'\alpha_j=\delta_{ij},\ i,j=1,2,\cdots,n,$$

即 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 $R^{(n)}$ 的标准正交基。

反之，由 $(\alpha_i,\alpha_j)=\alpha_i'\alpha_j=\delta_{ij}$；$i,j=1,2,\cdots,n$ 可知，$A'A=$

I_n，故A是正交阵。

有限n维欧氏空间V中任一向量α在标准正交基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的表示式设为

$$\alpha = \sum_{i=1}^{n} k_i \varepsilon_i,$$

则由ε_i的标准正交性可知

$$k_i = (\alpha, \varepsilon_i), \ i = 1, 2, \cdots, n,$$

称$k_1, k_2, \cdots, k_n$为α的正交坐标。

由于标准正交基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$的Gram矩阵是

$$A = ((\varepsilon_i, \varepsilon_j))_{n\times n} = (\delta_{ij})_{n\times n} = I_n.$$

设n维欧氏空间的向量α、β在$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的正交坐标分别为$x_1, x_2, \cdots, x_n$与$y_1, y_2, \cdots, y_n$，则

$$(\alpha, \beta) = x'Ay = x'y = \sum_{i=1}^{n} x_i y_i,$$

故在标准正交基下，两个向量内积的表示式具有特别简单的形状。

推论 2 设$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$是$n$维欧氏空间$V$的标准正交组，则必存在$\varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n$，使$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n$是$V$的标准正交基。

【证明】 因为V是n维线性空间，故必存在向量$\beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$，使$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$线性无关。由于$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$已是标准正交组，故只要对$\beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$用公式(1)的标准正交化方法，依次把$\alpha_{r+1}$写成$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \beta_{r+1}$的线性组合；把$\alpha_{r+2}$写成$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \alpha_{r+1}, \beta_{r+2}$的线性组合；…；把$\alpha_n$写成$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_{n-1}, \beta_n$的线性组合，即得正交向量组$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_n$，然后把它们单位化为$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n$，它们就是$V$的标准正交基。 证毕。

设Q_1是$n\times r$实阵，且满足：$Q_1'Q_1 = I_r$，则称Q_1为**列正交阵**。易知列正交阵Q_1的r个列向量构成一个标准正交组，因此列正交阵显然是列满秩阵。若把推论2用于n维欧氏空间$R^{(n)}$，便得到矩

阵论上的一个有用结论，即

推论 3 对任何 $n\times r$ 列正交阵 Q_1，必存在 $n\times(n-r)$ 列正交阵 Q_2，使 $Q=(Q_1, Q_2)$ 是 n 阶正交阵。

【证明】 把 Q_1 按列分块：$Q_1=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r)$，则 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r$ 是 $R^{(n)}$ 的一个标准正交组，故由推论 2，存在 $\varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n\in R^{(n)}$，使 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n$ 是 $R^{(n)}$ 的标准正交基，所以（由例 3），$Q=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n)=(Q_1, Q_2)$是正交阵，其中，$Q_2=(\varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n)$，由于 $\varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_n$ 显然是标准正交组，故 $Q_2'Q_2=I_{n-r}$，即 Q_2 是 $n\times(n-r)$列正交阵。 证毕。

例 5 设 x_i, y_i 全是实数，且 $\sum\limits_{i=1}^{n}x_i^2=\sum\limits_{i=1}^{n}y_i^2=1$，$\sum\limits_{i=1}^{n}x_iy_i=0$，求证：对任一 i，$1\leqslant i\leqslant n$，恒有：

$$x_i^2+y_i^2\leqslant 1. \tag{3}$$

【证明】 记 $\varepsilon_1=(x_1, x_2, \cdots, x_n)'$，$\varepsilon_2=(y_1, y_2, \cdots, y_n)'$，$Q_1=(\varepsilon_1, \varepsilon_2)$，则由假设可知，$Q_1$ 是 $n\times 2$ 列正交阵，由推论 3，存在 $n\times(n-2)$列正交阵 Q_2，使(Q_1, Q_2)为 n 阶正交阵，设 Q_2 的第 i 行是$(z_3^{(i)}, z_4^{(i)}, \cdots, z_n^{(i)})$，则由$(Q_1, Q_2)$的正交性显然可得

$$x_i^2+y_i^2+\sum_{j=3}^{n}(z_j^{(i)})^2=1,$$

证 $x_i^2+y_i^2\leqslant 1$。 证毕。

二、化二次型到主轴上去的另一方法

第七章中已用找镜象阵的方法找出了化实对称阵为对角阵的正交阵。本段将用另一方法找出这种正交阵。

先将实对称阵 A 的特征向量看作欧氏空间 $R^{(n)}$ 的向量（内积是 $\alpha'\beta$），则有下面的：

引理 属于实对称阵 A 的不同的特征值的特征向量必正交。

【证明】 设 λ, μ 是 A 的特征值，$\lambda\neq\mu$。又设 α 与 β 分别是属于 λ 与 μ 的特征向量，则由

$$\lambda\alpha'\beta=\lambda\beta'\alpha=\beta'A\alpha=\alpha'A\beta=\mu\alpha'\beta,$$

可得

$$(\lambda-\mu)(\alpha'\beta)=0,$$

由于 $\lambda\neq\mu$，所以 $\alpha'\beta=0$，即 α 与 β 正交。　　　　证毕。

引理 2　设 n 阶实对称阵 A 的特征值 λ 的重数是 k，则 A 有且只有 k 个属于 λ 的线性无关的特征向量。

【证明】　设 A 的其他特征值是 $\mu_{k+1},\mu_{k+2},\cdots,\mu_n,\lambda\neq\mu_i,i=k+1,k+2,\cdots,n$，则因 A(正交)相似于$[\lambda,\cdots,\lambda,\mu_{k+1},\mu_{k+2},\cdots,\mu_n]$，故 λI_n-A 相似于$[0,\cdots,0,\lambda-\mu_{k+1},\lambda-\mu_{k+2},\cdots,\lambda-\mu_n]$，于是 λI_n-A 的秩为 $n-k$，所以$(\lambda I_n-A)x=0$ 的基础解系有且只有 k 个线性无关的特征向量。　　　　证毕。

对于已给的 n 阶实对称阵 A，可由下面三个步骤找出正交阵 Q，使 $Q'AQ$ 为对角阵。

(i) 算出 A 的所有不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 以及它们的重数 $n_1,n_2,\cdots,n_s$，则 $\sum\limits_{i=1}^{s}n_i=n$；

(ii) 分别求出属于 λ_i 的极大线性无关向量，由引理 2，可设它们分别是：

$$\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in_i},\quad i=1,2,\cdots,s,$$

显然，

$$A\alpha_{ij}=\lambda_i\alpha_{ij},\ \ i=1,2,\cdots,s;\ \ j=1,2,\cdots,n_i;$$

(iii) 用 Gram-Schmidt 正交化方法，在欧氏空间 $R^{(n)}$ 中，把 $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in_i}$ 标准正交化为：

$$\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{in_i},\ i=1,2,\cdots,s$$

即

$$\varepsilon'_{ij}\varepsilon_{ik}=\delta_{jk},\tag{4}$$

因为 ε_{ij} 都是 $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in_i}$ 的线性组合，故

$$A\varepsilon_{ij}=\lambda_i\varepsilon_{ij},\ i=1,2,\cdots,s;\ \ j=1,2,\cdots,n_i.\tag{5}$$

这说明 $\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{in_i}$ 都是属于 λ_i 的特征向量，$i=1,2,\cdots,s$，且由引理 1 可知

$$\varepsilon'_{ij}\varepsilon_{kl}=0,\quad i\neq k.\tag{6}$$

作 n 阶阵：

$$Q=(\varepsilon_{11}, \cdots, \varepsilon_{1n_1}, \varepsilon_{21}, \cdots, \varepsilon_{2n_2}, \cdots, \varepsilon_{s1}, \cdots, \varepsilon_{sn_s}), \qquad (7)$$

则由(4)式与(6)式可知 Q 是正交阵，且由(5)式得到

$$AQ=Q[\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \cdots, \lambda_s I_{n_s}],$$

即

$$Q'AQ=[\lambda_1 I_{n1}, \lambda_2 I_{n_2}, \cdots, \lambda_s I_{n_s}],$$

故 Q 即为所求之正交阵。

例 6 仍以第七章 § 2 的例 2 为例. 因为方阵

$$A=\begin{pmatrix} 0 & 1 & 1 & - \\ 1 & 0 & -1 & \\ 1 & -1 & 0 & \\ -1 & 1 & 1 & \end{pmatrix}$$

的特征值是 1, 1, 1, −3, 故可算出：

属于 $\lambda=1$ 的三个线性无关的特征向量是：

$$\alpha_{11}=(1, 1, 0, 0)',$$
$$\alpha_{12}=(1, 0, 1, 0)',$$
$$\alpha_{13}=(-1, 0, 0, 1)'.$$

它们的标准正交化向量组是：

$$\varepsilon_{11}=\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right),$$

$$\varepsilon_{12}=\left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0\right)',$$

$$\varepsilon_{13}=\left(-\frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{3}{\sqrt{12}}\right)'.$$

而属于 $\lambda=-3$ 的一个特征向量是

$$\alpha_{21}=(1, -1, -1, 1),$$

它的单位化向量是

$$\varepsilon_{21}=\left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)',$$

故所求之正交阵为

$$Q=(\varepsilon_{11},\ \varepsilon_{12},\ \varepsilon_{13},\ \varepsilon_{21})=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{12}} & \frac{1}{2}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{-1}{2}\\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & \frac{-1}{2}\\ 0 & 0 & \frac{3}{\sqrt{12}} & \frac{1}{2}\end{pmatrix}.$$

必须说明，在一般情形下，仍以用镜象阵的方法求正交阵来得方便，只有在用镜象阵处理出现比较复杂的情形下，才用本段的方法去求正交阵 Q。

三、欧氏空间的同构

定义 设 σ 是欧氏空间 V 映到欧氏空间 V' 上的双射，如果对 V 中任意向量 α、β 以及任意实数 k 恒有

(i) $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$；

(ii) $\sigma(k\alpha)=k\sigma(\alpha)$；

(iii) $(\sigma(\alpha),\ \sigma(\beta))=(\alpha,\ \beta)$。

则称 σ 是欧氏空间 V 映到 V' 上的**同构映射**。若有这样的 σ，则称 V 与 V' 是**同构的**。仍以 $V\cong V'$ 记之。

由欧氏空间的同构映射定义的(i)、(ii)可知，σ 首先是线性空间的同构，同时由(iii)可知，σ 保持内积不变。

定理 2.2 两个有限维欧氏空间同构的充要条件是它们的维数相同。

【证明】 必要性是明显的，因为欧氏空间的同构首先是线性空间的同构，故由第九章定理 6.1 可知，这两个欧氏空间的维数相同。

充分性．设 V、V' 是两个 n 维欧氏空间，$\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n$ 与 $\varepsilon_1',\ \varepsilon_2',\ \cdots,\ \varepsilon_n'$ 分别是 V 与 V' 的标准正交基，任取 $\alpha\in V$，则

$$\alpha=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n,$$

作映射：

σ: $\quad \alpha=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n \to k_1\varepsilon_1'+k_2\varepsilon_2'+\cdots+k_n\varepsilon_n'=\sigma(\alpha)$,

第九章定理 6.1 已证得：若

σ: $\quad \beta=l_1\varepsilon_1+l_2\varepsilon_2+\cdots+l_n\varepsilon_n \to l_1\varepsilon_1'+l_2\varepsilon_2'+\cdots+l_n\varepsilon_n'=\sigma(\beta)$,

则

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),$$
$$\sigma(k\alpha)=k\sigma(\alpha).$$

又因

$$(\alpha,\ \beta)=\sum_{i=1}^{n}k_il_i,\quad (\sigma(\alpha),\ \sigma(\beta))=\sum_{i=1}^{n}k_il_i,$$

所以$(\sigma(\alpha),\ \sigma(\beta))=(\alpha,\ \beta)$，这说明$\sigma$使$V$与$V'$同构。　　证毕。

定理 2.2 有些有趣的应用，它虽然说的是有限维欧氏空间，但对无限维欧氏空间中只涉及有限个向量的一些问题仍能应用它。例如，对任何欧氏空间中的任意三个向量α、β、γ，作线性包$L(\alpha,\ \beta,\ \gamma)$，则$L(\alpha,\ \beta,\ \gamma)$是$V$的子空间，且它的维数不超过 3，所以由定理 2.2 可知，$L(\alpha,\ \beta,\ \gamma)$与通常的三维(几何)欧氏空间的一个子空间同构。因此，要找有关这三个向量α、β、γ的任何结论，只要在通常的解析几何中去找就可以了(指涉及加法、数乘与内积这三种运算的结论而言)。例如，在解析几何中，任意两个向量的数量积(内积)不大于两个向量的长度的乘积，故按照定理 2.2 欧氏空间同构的观点可知，对任何欧氏空间中的向量$\alpha,\ \beta$，必有：

$$|(\alpha,\ \beta)|\leqslant(\alpha,\ \alpha)^{1/2}(\beta,\ \beta)^{1/2}$$

(因$\alpha,\ \beta\in L(\alpha,\ \beta,\ \gamma)$，而$L(\alpha,\ \beta,\ \gamma)$与三维空间的子空间同构)。这就是 §1 中已经证明过的 Cauchy–Буняковский 不等式。

又如，平面几何中有"三角形两边之和不小于第三边"。用定理 2.2 同构的观点，即可知对任意欧氏空间中的向量α、β，恒有

$$\sqrt{(\alpha+\beta,\ \alpha+\beta)}\leqslant\sqrt{(\alpha,\ \alpha)}+\sqrt{(\beta,\ \beta)},$$

即

$$\|\alpha+\beta\|\leqslant\|\alpha\|+\|\beta\|.$$

由此可见定理 2.2 是个十分有用的结果，它指明了同构的意义与作用所在。

习　题

1. 将多项式空间 $R[x]$ 的线性无关向量 x, x^3, x^5, x^7 标准正交化，内积是 $(f, g)=\int_0^1 f(x)g(x)\mathrm{d}x$。

2. 把 $R^{(2)}$ 的向量 $(1, 1)'$，$(-1, 1)'$ 标准正交化，内积是 $(\alpha, \beta)=\alpha' A\beta$，其中 $A=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$。

3. 把 Gram-Schmidt 方法用于带有内积 $\alpha'\beta$ 的 n 维列向量空间 $R^{(n)}$，证明：任何非异实方阵必可分解为一个正交阵与一个非异上三角实方阵的乘积。

4. 运用第二段的方法求正交阵 Q，使 $Q'AQ$ 为对角阵。

(i) $A=\begin{pmatrix}3 & -2 & -4\\ -2 & 6 & -2\\ -4 & -2 & 3\end{pmatrix}$；(ii) $A=\begin{pmatrix}-1 & -3 & 3 & -3\\ -3 & -1 & -3 & 3\\ 3 & -3 & -1 & -3\\ -3 & 3 & -3 & -1\end{pmatrix}$。

5. 求证：n 维欧氏空间的标准正交基 $e_1, e_2, \cdots, e_n$ 到标准正交基 $e_1, e_2, \cdots, e_n$ 的过渡矩阵 P 是正交阵。

6. 求证：秩为 r 的 n 阶实方阵必有分解式：$A=(Q, 0)B$，其中 Q 是 $n\times r$ 列正交阵，B 是 n 阶非异阵。

7. 用欧氏空间同构的观点具体证明：

(i) 任何欧氏空间中的两个向量 α、β 必满足不等式：

$$(\alpha, \beta)^2\leqslant(\alpha, \alpha)(\beta, \beta),$$

(ii) 任何欧氏空间中三个向量 α、β、γ 必满足不等式：

$$\rho(\alpha, \gamma)\leqslant\rho(\alpha, \beta)+\rho(\beta, \gamma),$$

此处 $\rho(\alpha, \beta)$ 表示 α 与 β 的距离；

(iii) 对 $R[a, b]$ 中任意两个向量 $f(x)$，$g(x)$，恒有：

$$\left(\int_a^b (f(x)+g(x))^2\mathrm{d}x\right)^{1/2}\leqslant\left(\int_a^b f(x)^2\mathrm{d}x\right)^{1/2}+\left(\int_a^b g(x)^2\mathrm{d}x\right)^{1/2}.$$

§3　共轭变换与自共轭变换、正交变换

一、共轭变换与自共轭变换

定义　设 σ 是 n 维欧氏空间 V 的任一线性变换，σ 在 V 的标

准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵是 A，则称 A' 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下所对应的线性变换为(σ 的)**共轭(线性)变换**，以 σ' 记之。有时也称 σ' 为(σ 的)**转置变换**。

易证 σ' 是由 σ 唯一决定的。由共轭变换的定义易知

$$\sigma'(\varepsilon_k)=\sum_{j=1}^{n} a_{jk}\varepsilon_j, \quad k=1, 2, \cdots, n, \tag{1}$$

且 σ' 与 σ 有如下关系式：

命题 1 设 σ' 是 σ 的共轭变换，则

$$(\sigma(\varepsilon_i), \varepsilon_k)=(\varepsilon_i, \sigma'(\varepsilon_k))。\tag{2}$$

【证明】 因为

$$(\sigma(\varepsilon_i), \varepsilon_k)=\left(\sum_{j=1}^{n} a_{ij}\varepsilon_j, \varepsilon_k\right)=\sum_{j=1}^{n} a_{ij}(\varepsilon_j, \varepsilon_k)$$

$$=\sum_{j=1}^{n} a_{ij}\delta_{jk}=a_{ik},$$

$$(\varepsilon_i, \sigma'(\varepsilon_k))=\left(\varepsilon_i, \sum_{j=1}^{n} a_{jk}\varepsilon_j\right)=\sum_{j=1}^{n} a_{jk}(\varepsilon_i, \varepsilon_j)$$

$$=\sum a_{jk}\delta_{ij}=a_{ik},$$

故(2)式是正确的。 证毕。

定理 3.1 设 σ 是 n 维欧氏空间 V 的线性变换，则 V 的线性变换 τ 是 σ 的共轭变换的充要条件是，对 V 中任何向量 α、β，恒有：

$$(\sigma(\alpha), \beta)=(\alpha, \tau(\beta))。\tag{3}$$

【证明】 取 V 的标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，则 $\alpha=\sum_{i=1}^{n}k_i\varepsilon_i$，$\beta=\sum_{j=1}^{n}l_j\varepsilon_j$。

必要性。如果 $\tau=\sigma'$，则由命题 1，

$$(\sigma(\alpha), \beta)=\left(\sigma\left(\sum_{i=1}^{n}k_i\varepsilon_i\right), \sum_{j=1}^{n}l_j\varepsilon_j\right)$$

$$=\sum_{i,j=1}^{n} k_il_j(\sigma(\varepsilon_i), \varepsilon_j)$$

$$= \sum_{i,j=1}^{n} k_i l_j(\varepsilon_i, \sigma'(\varepsilon_j))$$

$$= \left(\sum_{i=1}^{n} k_i\varepsilon_i, \sigma\left(\sum_{j=1}^{n} l_j\varepsilon_j\right)\right) = (\alpha, \sigma'(\beta)).$$

充分性. 因为对 σ 的共轭变换 σ'，上面的必要性已证明必有：

$$(\sigma(\alpha), \beta) = (\alpha, \sigma'(\beta)), \tag{4}$$

今由充分条件的假设，成立(3)式，故由(3)式与(4)式即得

$$(\alpha, \sigma'(\beta)) = (\alpha, \tau(\beta)),$$

或者

$$(\alpha, (\sigma' - \tau)\beta) = 0.$$

由于 α、β 是任意的，故特别取 $\beta = \varepsilon_j$，$(\sigma' - \tau)\varepsilon_j = \alpha$，故得

$$(\sigma' - \tau)\varepsilon_j = 0, \ j = 1, 2, \cdots, n,$$

所以 $\sigma' - \tau = 0^*$，即 $\sigma' = \tau$。 **证毕.**

定义 设 σ 是 n 维欧氏空间的线性变换，如果 $\sigma' = \sigma$，则称 σ 为**自共轭变换**或**对称变换**。

定理 3.2 设 σ 是 n 维欧氏空间 V 的线性变换，则 σ 是自共轭变换的充要条件是：

(i) σ 在 V 的任一标准正交基下的矩阵是实对称阵，或者；

(ii) 对任何 $\alpha, \beta \in V$，恒有

$$(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta)). \tag{5}$$

【证明】 由定理 3.1 可知第(ii)个结论是显然的，今证(i)。设 σ 在 V 的标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵为 A，则

η： $$\sigma \to A,$$

又

η： $$\sigma' \to A',$$

但 $\sigma = \sigma'$，故由 η 的单值性，$A = A'$。这说明 A 是实对称阵。又设 σ 在标准正交基 $e_1, e_2, \cdots, e_n$ 下的阵为 B，则 $B = P^{-1}AP$，此处 P 是 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $e_1, e_2, \cdots, e_n$ 的过渡矩阵，但 P 是正交阵（见上节的习题第 5 题），所以 $P' = P^{-1}$，于是 $B = P'AP$ 仍是实对称阵。

反之，设 σ 在 V 的任一么正基下的阵为实对称阵 B，则由 $\sigma \to$

B, $\sigma\to B'$ 以及这个映射的单射性质可知,由于 $B=B'$, 故 $\sigma=\sigma'$。这说明 σ 是自共轭变换。 证毕。

由于 n 维欧氏空间的自共轭变换与 n 阶实对称阵一一对应,这就使我们能应用由解析方法得到的实对称阵的一系列结论去获得 n 维欧氏空间的自共轭变换的一系列有趣的"几何"性质。今略举例于下:

例 1 n 维欧氏空间 V 的自共轭变换 σ 必有 n 个两两正交的特征向量(称 σ 具有完全的正交特征向量系)。

【证明】 取定 V 的标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, 设 σ 在这个基下的矩阵为 A, 则由定理 3.2 可知 A 是实对称阵,且

η: $$\sigma\to A$$

是一双射。由于 A 是实对称阵,故存在正交阵 Q, 使

$$Q'AQ=[\lambda_1, \lambda_2, \cdots, \lambda_n], \tag{6}$$

此处 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的特征值.若把 Q 按列分块: $Q=(q_1, q_2, \cdots, q_n)$, 则 $q_1, q_2, \cdots, q_n$ 是 $R^{(n)}$ 的标准正交基, 即 $q_i'q_j=\delta_{ij}$, $i, j=1, 2, \cdots, n$。并且由(6)式可得

$$Aq_i=\lambda_i q_i, \quad i=1, 2, \cdots, n, \tag{7}$$

又在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下,设

δ: $$q_i\to\alpha_i, \ i=1, 2, \cdots, n,$$

则 $\alpha_i\in V$, 且

δ: $$\lambda_i q_i\to\lambda_i\alpha_i, \quad i=1, 2, \cdots, n,$$

δ: $$Aq_i\to\sigma(\alpha_i), \quad i=1, 2, \cdots, n$$

(见第十章定理 2.1)。故由(7)式及 δ 的单值性可知

$$\sigma(\alpha_i)=\lambda_i\alpha_i, \tag{8}$$

因为 $R^{(n)}$ 与 V 是同构的(欧氏空间的同构),故

$$(\alpha_i, \alpha_j)=\alpha_i'\alpha_j=q_i'q_j=\delta_{ij}, \quad i, j=1, 2, \cdots, n。 \tag{9}$$

(8)与(9)说明 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 σ 的特征向量,且两两正交。证毕。

例 1 说明, n 维欧氏空间中必有一个标准正交基作为空间的自共轭变换的特征向量组。在例 1 的证明过程中还附带证明了 λ_i 是 σ 的特征值,由此可知, n 维欧氏空间的自共轭变换的特征值全

是实数，且σ有关于V的一维不变子空间。

二、正交变换

定义 设σ为n维欧氏空间V的非异线性变换，如果$\sigma^{-1}=\sigma'$，则称σ为**正交变换**。

定理3.3 设σ为n维欧氏空间V的线性变换，则下列五个条件是等价的

(i) σ是正交变换；

(ii) σ是V映到自身上的同构映射，即σ保持任意两个向量的内积不变：$(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta)$；

(iii) σ保持向量的长度不变；

(iv) σ把标准正交基变为标准正交基；

(v) σ在任何标准正交基下的矩阵是正交阵。

【证明】 采用循环证法：(i)→(ii)→(iii)→(iv)→(v)→(i)。

(i)→(ii)：由于$\sigma^{-1}=\sigma'$，故$\sigma'\sigma=1^*$，对任何$\alpha,\beta\in V$，

$$(\sigma(\alpha),\sigma(\beta))=(\alpha,\sigma'(\sigma(\beta))=(\alpha,\beta). \tag{10}$$

(ii)→(iii)：取$\alpha=\beta$，于是$(\sigma(\alpha),\sigma(\alpha))=(\alpha,\alpha)$，即$\|\sigma(\alpha)\|=\|\alpha\|$。

(iii)→(iv)：取V的任一标准正交基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，则

$$(\sigma(\varepsilon_i),\sigma(\varepsilon_i))=(\varepsilon_i,\varepsilon_i)=1, \tag{11}$$

另外，当$i\neq j$时，由

$$(\sigma(\varepsilon_i+\varepsilon_j),\sigma(\varepsilon_i+\varepsilon_j))=(\varepsilon_i+\varepsilon_j,\varepsilon_i+\varepsilon_j),$$

可得

$$\begin{aligned}&2(\sigma(\varepsilon_i),\sigma(\varepsilon_j))+(\sigma(\varepsilon_i),\sigma(\varepsilon_i))+(\sigma(\varepsilon_j),\sigma(\varepsilon_j))\\&=2(\varepsilon_i,\varepsilon_j)+(\varepsilon_i,\varepsilon_i)+(\varepsilon_j,\varepsilon_j),\end{aligned}$$

故由(11)式可知，当$i\neq j$时上式化为

$$(\sigma(\varepsilon_i),\sigma(\varepsilon_j))=(\varepsilon_i,\varepsilon_j)=0,$$

所以$(\sigma(\varepsilon_i),\sigma(\varepsilon_j))=\delta_{ij}$，$i,j=1,2,\cdots,n$。即$\sigma(\varepsilon_1),\sigma(\varepsilon_2),\cdots,\sigma(\varepsilon_n)$是$V$的标准正交基。

(iv)→(v)：设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是$V$的标准正交基。

$$\sigma(\varepsilon_i)=\sum_{j=1}^{n}a_{ij}\varepsilon_j,\quad i=1,2,\cdots,n,$$

则由(iv)可知，

$$\delta_{ik}=(\sigma(\varepsilon_i),\sigma(\varepsilon_k))=\left(\sum_{j=1}^{n}a_{ij}\varepsilon_j,\sum_{s=1}^{n}a_{ks}\varepsilon_s\right)$$

$$=\sum_{j,s=1}^{n}a_{ij}a_{ks}(\varepsilon_j,\varepsilon_s)=\sum_{j=1}^{n}a_{ij}a_{kj},\ i,k=1,2,\cdots,n,$$

故方阵 $A=(a_{ij})_{n\times n}$ 是正交阵。

(v)→(i)：因 σ 在任一标准正交基下的矩阵 A 是正交阵，$\sigma\to A$，故 σ 为非异线性变换，因而 $\sigma^{-1}\to A^{-1}$，又 $\sigma'\to A'=A^{-1}$，故由映射的单射性可知，$\sigma^{-1}=\sigma'$，即 σ 是正交变换。　　证毕。

习　题

1. 设 σ、τ 是 n 维欧氏空间的线性变换，求证：

(i) $(\sigma+\tau)'=\sigma'+\tau'$；(ii) $(\sigma\tau)'=\tau'\sigma'$；(iii) $(\sigma')'=\sigma$。

2. 求证：n 维欧氏空间 V 必可分解为 n 个关于 V 的自共轭变换 σ 的一维不变子空间的直接和。

3. 求证：n 维欧氏空间 V 必可分解为有限个关于 V 的正交变换 σ 的一维不变子空间与二维不变子空间的直接和。

4. n 维欧氏空间的正交变换是否必有两两正交的完全特征向量系？

5. 设 σ 是 n 维欧氏空间 V 的线性变换，求证：

$$\langle\alpha,\beta\rangle=\langle\sigma(\alpha),\sigma(\beta)\rangle,\ \forall\alpha,\beta\in V,$$

即正交变换保持两向量间的夹角不变。反之，若 n 维欧氏空间的线性变换 σ 保持任意两个向量的交角不变，问 σ 是否正交变换？

6. 求证：n 维欧氏空间 V 的线性变换 σ 是正交变换的充要条件是，对任何 α、$\beta\in V$，恒有

$$\rho(\alpha,\beta)=\rho(\sigma(\alpha),\sigma(\beta)),$$

即正交变换保持两向量之间的距离不变。

7. 设 u 是 n 维欧氏空间 V 的单位向量，定义

$$\sigma(\alpha)=\alpha-2(u,\alpha)u,$$

称 σ 为**镜象变换**或**镜面反射变换**。求证：

(i) σ 是正交、自共轭变换；

(ii) σ 有一个特征值是 -1，有 $n-1$ 个特征值是 1；

(iii) 任何正交变换必可分解为有限个镜象变换的乘积。

§4 正射影、最小平方偏差问题

一、正交补空间

设W是欧氏空间V的任意一个非平凡子空间，则W对于V的内积仍是一个欧氏空间．取V中向量α，如果α与W中任一向量正交，则称α **垂直于** W，以$\alpha\perp W$记之。又记

$$W^{\perp}=\{\alpha|\alpha\perp W,\ \alpha\in V\}.$$

命题1 $W^{\perp}$是V的(欧氏)子空间。

【证明】 任取$\alpha,\beta\in W^{\perp}$，$\delta$是$W$中任一向量，则

$$(k_1\alpha+k_2\beta,\ \delta)=k_1(\alpha,\ \delta)+k_2(\beta,\ \delta)=0,$$

于是$k_1\alpha+k_2\beta\in W^{\perp}$，即$W^{\perp}$是$V$的(欧氏)子空间。 证毕。

定理4.1 对任何欧氏空间V的有限维非平凡子空间W，必存在唯一的一个子空间$W^{\perp}$，使

$$V=W\oplus W^{\perp}. \tag{1}$$

【证明】 设W的维数是k，因为W也是欧氏空间，故可设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_k$是$W$的标准正交基。任取$\alpha\in V$，作向量

$$\beta=\sum_{i=1}^{k}(\alpha,\ \varepsilon_i)\varepsilon_i,$$

则$\beta\in W$，令

$$\delta=\alpha-\beta=\alpha-\sum_{i=1}^{k}(\alpha,\ \varepsilon_i)\varepsilon_i, \tag{2}$$

对任一j，$1\leqslant j\leqslant k$，由于

$$\begin{aligned}(\delta,\ \varepsilon_j)&=\left(\alpha-\sum_{i=1}^{k}(\alpha,\ \varepsilon_i)\varepsilon_i,\ \varepsilon_j\right)\\&=(\alpha,\varepsilon_j)-\sum_{i=1}^{k}(\alpha,\ \varepsilon_i)(\varepsilon_i,\ \varepsilon_j)\\&=(\alpha,\ \varepsilon_j)-\sum_{i=1}^{k}(\alpha,\ \varepsilon_i)\delta_{ij}=0,\end{aligned}$$

这说明$\delta\perp\varepsilon_j$，$j=1,2,\cdots,k$，所以$\delta\perp W$，于是$\delta\in W^{\perp}$，故(2)式可写成：$\alpha=\beta+\delta$，即$V\subseteq W+W^{\perp}$，显然$V\supseteq W+W^{\perp}$，故得

$$V=W+W^{\perp}. \tag{3}$$

今证上述子空间之和为直接和。因为对 $W\cap W^{\perp}$ 的向量 α，由于 $\alpha\in W$，$\alpha\in W^{\perp}$，故 $(\alpha,\alpha)=0$，即 $\alpha=\theta$，所以(3)是直接和：

$$V=W\oplus W^{\perp}.$$

今证唯一性，设存在 V 的子空间 U，使

$$V=W\oplus U, \tag{4}$$

且 U 中任一向量 $\delta_1\perp W^{\perp}$，则必 $U=W^{\perp}$，这因为：由 $W^{\perp}$ 的定义，显然有 $U\subseteq W^{\perp}$，今任取 $\delta\in W^{\perp}$，则由(4)式，

$$\delta=\beta+\delta_1,\ \beta\in W,\ \delta_1\in U,$$

因为 $\delta\perp\beta$，$\delta_1\perp\beta$，故

$$\theta=(\delta,\beta)=(\beta+\delta_1,\beta)=(\beta,\beta),$$

即 $\beta=\theta$，所以 $\delta=\delta_1\in U$，也即 $W^{\perp}\subseteq U$，所以 $W^{\perp}=U$。　　证毕。

必须注意：当 V 是有限维欧氏空间时，唯一性是显然的，因为 $U\subseteq W^{\perp}$，且 $\mathrm{d}(U)=\mathrm{d}(W^{\perp})$，故 $U=W^{\perp}$。但当 V 是无限维欧氏空间时，唯一性并不是显然的。

定义　称 $W^{\perp}$ 为W的**正交补空间**。

当 V 是有限维欧氏空间时，由于

$$V=W\oplus W^{\perp}=W^{\perp}\oplus W,$$

$$V=W^{\perp}\oplus(W^{\perp})^{\perp},$$

故由定理 4.1 的唯一性，即知

$$(W^{\perp})^{\perp}=W, \tag{5}$$

故W与 $W^{\perp}$ 互为正交补(空间)。

二、正射影

设W是欧氏空间 V 的一个有限维非平凡子空间。取 V 中向量 α，今要求在W中找一个向量 α_W，使 $(\alpha-\alpha_W)\perp W$。这个问题在有限维欧氏空间、特别是三维(欧氏)几何空间有明显的几何意义，那就是，α_W 是“点” α 在平面W上的正射影，而 $\alpha-\alpha_W$ 就是过“点” α 到W的垂线。它的长度显然是“点” α

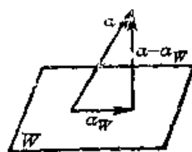

图 1

到W的最短距离(见图 1)。

命题 2 如果W是欧氏空间V的任一有限非平凡子空间，则对任一$\alpha\in V$，必存在唯一的一个$\alpha_W\in W$，使$(\alpha-\alpha_W)\perp W$。

【证明】 由定理 4.1，对任一$\alpha\in V$，必有分解式：

$$\alpha=\alpha_W+\alpha_{W^\perp},$$

其中$\alpha_W\in W$，$\alpha_{W^\perp}\in W^\perp$，且$\alpha_W$与$\alpha_{W^\perp}$由$\alpha$唯一决定。因为$\alpha_W\perp\alpha_{W^\perp}$，而

$$\alpha_{W^\perp}=\alpha-\alpha_W,$$

于是$(\alpha-\alpha_W)\perp\alpha_W$，故$\alpha_W$即为所求。 证毕。

命题 3 $\rho(\alpha,\alpha_W)=\|\alpha-\alpha_W\|$是$\alpha$到$W$的"最短距离"，即对$W$中另一向量$\beta\neq\alpha_W$，必有

$$\rho(\alpha,\alpha_W)<\rho(\alpha,\beta)。\tag{6}$$

【证明】 因$\beta\neq\alpha_W$，故$\alpha_W-\beta\neq\theta$，且$\alpha_W-\beta\in W$，所以$(\alpha_W-\beta)\perp(\alpha-\alpha_W)$，由商高定理可知

$$\|\alpha-\beta\|^2=\|\alpha-\alpha_W\|^2+\|\alpha_W-\beta\|^2>\|\alpha-\alpha_W\|^2$$

(因为$\alpha_W-\beta\neq\theta$)，所以$\rho(\alpha,\beta)>\rho(\alpha,\alpha_W)$。 证毕。

由于命题 2 与命题 3，我们看到α_W具有通常三维几何空间的向量的正射影完全相同的性质，故引进

定义 设W是欧氏空间V的有限维非平凡子空间，$\alpha\in W$，且$\alpha=\alpha_W+\alpha_{W^\perp}$，其中$\alpha_W\in W$，$\alpha_{W^\perp}\in W^\perp$，则称$\alpha_W$为$\alpha$在$W$上的**正射影**。

今来具体求出α在W上的正射影α_W，我们有

定理 4.2 设$\alpha_1,\alpha_2,\cdots,\alpha_k$是$W$的任意一个基，则向量：$x_1\alpha_1+x_2\alpha_2+\cdots+x_k\alpha_k$是$\alpha$在$W$上正射影的充要条件是，$x_1,x_2,\cdots,x_k$为下列线性方程组的解：

$$\sum_{j=1}^{n}(\alpha_i,\alpha_j)x_j=(\alpha,\alpha_i),\ i=1,2,\cdots,k,\tag{7}$$

也即$x=(x_1,x_2,\cdots,x_k)'$为矩阵方程

$$G(\alpha_1,\alpha_2,\cdots,\alpha_k)x=b\tag{8}$$

的解，而$b=((\alpha,\alpha_1),(\alpha,\alpha_2),\cdots,(\alpha,\alpha_k))'$。

【证明】 如果 $\alpha_W=x_1\alpha_1+x_2\alpha_2+\cdots+x_k\alpha_k$ 是 α 在W上的正射影．则因$(\alpha-\alpha_W)\perp W$，故对任一 i，$1\leqslant i\leqslant k$，恒有

$$0=\left(\alpha-\sum_{j=1}^{k}x_j\alpha_j,\ \alpha_i\right)=(\alpha,\ \alpha_i)-\sum_{j=1}^{k}x_j(\alpha_i,\ \alpha_j),\ i=1,\ 2,\ \cdots,\ k;$$

故得(7)式。

反之，由于 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 线性无关，故 $|G(\alpha_1, \alpha_2, \cdots, \alpha_k)|\neq 0$，因此(7)的解是唯一的。于是把上面必要性的证明过程依次倒推回去，所求出唯一的 $\sum_{j=1}^{k}x_j\alpha_j$ 就是 α 在W上的正射影。

证毕。

如果在定理 4.2 中取W的基为标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k$，则$(\varepsilon_i, \varepsilon_j)=\delta_{ij}$，故由(7)式即得

$$x_i=(\alpha,\ \varepsilon_i),\ i=1,\ 2,\ \cdots,\ k, \tag{9}$$

即 x_i 是 α_W 在 ε_i 上的坐标。这与通常三维几何空间的结论是一致的。所以当取W的标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k$ 时，我们立刻可以写出 α 在W上的正射影 α_W 的明确表达式。

$$\alpha_W=\sum_{i=1}^{k}(\alpha,\ \varepsilon_i)\varepsilon_i. \tag{10}$$

(10)式给了定理 4.1 的证明过程以更明确的几何意义及证明思路。

三、最小平方偏差问题

某些实际问题及理论问题常归结为：对欧氏空间V的向量 α，要找V的有限维非平凡子空间W中的向量 β，使$(\rho(\alpha, \beta))^2=\|\alpha-\beta\|^2$ 最小。这个问题就称为**最小平方偏差问题**。本节第二段已证过，只要取 $\beta=\alpha_W$ 就可解决这一问题，所以最小平方偏差问题实际上就是要求 α 在W上的正射影。今举出它的两个应用：

(i) 最小二乘法

在生产实践与科学试验中遇到的某些问题所归结出来的数学模型，往往是一个没有解的 $m\times n$ 线性代数方程组（称为矛盾方程组），讨论这一类方程组具有较大的实际意义，矛盾方程组常常是

这样归结出来的：根据某些科学规律或预测，量 b（实数）与量 a_1, a_2, …, a_n（都是实数）之间存在着线性关系式

$$b=x_1a_1+x_2a_2+\cdots+x_na_n \tag{11}$$

这个关系式可能是（相对地）精确的，也可能是近似的反映，关系式中的 $x_1, x_2, \cdots, x_n$ 是未知的常系数（实数），它们可以由对 $a_1, a_2, \cdots, a_n$; b 进行一系列测试决定出来。设对 $a_1, a_2, \cdots, a_n$; b 进行了 m 次测试，测试出来的数据是

$$a_{i1}, a_{i2}, \cdots, a_{in}; b_i \quad (i=1, 2, \cdots, m)$$

则由关系式(11)可知成立下面各式：

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i, \quad i=1, 2, \cdots, m, \tag{12}$$

或者

$$Ax=b, \tag{13}$$

其中 $A=(a_{ij})_{m\times n}$ 是(12)的系数矩阵，$x=(x_1, x_2, \cdots, x_n)'$，$b=(b_1, b_2, \cdots, b_m)'$。现在的问题是要求出 x。由于对关系式(11)的认识的局限性（可能没有归结好），以及测试仪器产生的测量误差，因此方程组(12)（或(13)）总是一个矛盾方程组，要谈它的精确解是不可能的，反而是无意义的。因此只能这样找 x，使 Ax 尽可能接近 b。用式子表达，也就是要求 x 的"最佳近似值"$x=(x_1, x_2, \cdots, x_n)$，使

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n}a_{ij}x_j-b_i\right)^2 \tag{14}$$

（总体）的值最小。在这样意义下得到的解叫作矛盾方程组 $Ax=b$ **的极小二乘解。**

一般说来，m 恒大于 n，即测量次数恒多于量 $a_1, a_2, \cdots, a_n$ 的个数，这样可以提高 x 的精确度，例如在 1978 年，我国搞的一个"大地测量"问题所归结出来的线性方程组就是一个 17 万个方程、13 万个未知量的矛盾方程组。

下面找出这种 x。记 $A=(\alpha_1, \alpha_2, \cdots, \alpha_n)$，其中 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 都是 m 维列向量，则上述问题（(14)式）就归结为：找 $x=(x_1, x_2, \cdots, x_n)'$，使

$$\|b-(x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n)\|^2$$

最小，或者

$$\|b-Ax\|^2$$

最小。也就是说，要求欧氏空间 $R^{(m)}$ 中向量 b 到子空间 $W=L(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 的最短距离。这个最小平方偏差问题就称为**最小二乘方**问题，解决这一问题的方法就称作**最小二乘法**。

由上一段可知，我们的目的是找 b 在W上的正射影 b_W。而由定理 4.2 可知，$b_W=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$ 的 $x_1, x_2, \cdots, x_n$ 必可由下列线性方程组解出：

$$G(\alpha_1, \alpha_2, \cdots, \alpha_n)x=\begin{pmatrix}(\alpha_1, b)\\(\alpha_2, b)\\\vdots\\(\alpha_n, b)\end{pmatrix}=\begin{pmatrix}\alpha_1'b\\\alpha_2'b\\\vdots\\\alpha_n'b\end{pmatrix}=A'b, \qquad (15)$$

而

$$G(\alpha_1, \alpha_2, \cdots, \alpha_n)=\begin{pmatrix}\alpha_1'\alpha_1\cdots\alpha_1'\alpha_n\\\cdots\cdots\cdots\cdots\\\alpha_n'\alpha_1\cdots\alpha_n'\alpha_n\end{pmatrix}=A'A,$$

于是(15)式化为

$$A'Ax=A'b, \qquad (16)$$

即正射影 $b_W=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$ 的 $x=(x_1, x_2, \cdots, x_n)'$ 可由方程组(16)解出，这说明矛盾方程组 $Ax=b$ 的极小二乘解可以在相容方程组(16)的诸解中找出。特别当A是列满秩阵时，$A'A$ 是非异阵，此时(16)只有唯一解：

$$x=(A'A)^{-1}A'b, \qquad (17)$$

这个 x 自然就是矛盾方程组 $Ax=b$ 的极小二乘解。

如果称 $Ax=b$ 的所有极小二乘解中长度最小者为**最小二乘解**，则解(17)也是最小二乘解。

例 1 矛盾方程组：

$$\begin{cases}2x_1-x_2=3\\3x_1+x_2=4\\x_1+2x_2=2\end{cases}$$

的极小二乘解(恰好是最小二乘解)是

$$x=(A'A)^{-1}(A'b)=\begin{pmatrix}14 & 3\\ 3 & 6\end{pmatrix}^{-1}\begin{pmatrix}20\\ 5\end{pmatrix}=\begin{pmatrix}7/5\\ 2/15\end{pmatrix}.$$

(ii) 函数的逼近问题

设 $f(x)$是$[0, 2\pi]$上的连续实变实值函数,在某些逼近理论问题中,要求用一些已知的三角多项式"逼近"它,明确地说,就是要在所有的"三角多项式":

$$p(x)=\frac{a_0}{2}+a_1\cos x+b_1\sin x+\cdots+ak\cos kx+b_k\sin kx \tag{18}$$

中,找出某一个 $p(x)$,使 $\int_0^{2\pi}(f(x)-p(x))^2dx$ 最小,其中 a_0, a_i, b_i 都是实数, $i=1, 2, \cdots, k$.

这也是一个最小平方偏差问题,可用下法解决.对连续函数空间 $R[0, 2\pi]$,定义内积 $(f, g)=\int_0^{2\pi}f(x)g(x)dx$,则 $R[0, 2\pi]$是一个无限维欧氏空间,又第九章已证 $1, \cos x, \cdots, \cos kx, \sin x, \cdots, \sin kx$ 是 $R[0, 2\pi]$中 $2k+1$ 个线性无关的向量,作子空间:

$$W=L(1, \cos x, \sin x, \cdots, \cos kx, \sin kx),$$

则W是 $R[0, 2\pi]$的 $2k+1$ 维子空间。且W中任一元素都是形如(18)式的三角多项式 $p(x)$。故上面的问题就化为:求"点"$f(x)$到子空间W的最短距离,即找 $f(x)$在W上的正射影$(f(x))_W$.

由于 $1, \cos x, \sin x, \cdots, \cos kx, \sin kx$ 就是W的一个正交基(§2的例2),把这个正交基标准化,即得W的下列标准正交基:

$$\varepsilon_0=\frac{1}{\sqrt{2\pi}},\quad \varepsilon_1=\frac{\cos x}{\sqrt{\pi}},\quad \varepsilon_2=\frac{\sin x}{\sqrt{\pi}},\quad \cdots,\quad \varepsilon_{2k-1}=\frac{\cos kx}{\sqrt{\pi}},$$

$$\varepsilon_{2k}=\frac{\sin kx}{\sqrt{\pi}},$$

故由上节第(10)式可知, $f(x)$在W上的正射影: $(f(x))_W=p(x)$应是:

$$p(x)=\sum_{j=1}^{k}k_j\varepsilon_j, \tag{19}$$

而

$$x_j=(f,\ \varepsilon_j)=\int_0^{2\pi} f(x)\varepsilon_j(x)\mathrm{d}x,\quad j=0,1,\cdots,2k,$$

于是

$$x_0=\frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(x)\mathrm{d}x;$$

$$x_{2l-1}=\frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\cos lx\mathrm{d}x;$$

$$x_{2l}=\frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\sin lx\mathrm{d}x;$$

$l=1,2,\cdots,k$。由于(19)式可写成

$$p(x)=\frac{x_0}{\sqrt{2\pi}}+\sum_{l=1}^{k}\left(\frac{x_{2l-1}}{\sqrt{\pi}}\cos lx+\frac{x_{2l}}{\sqrt{\pi}}\sin lx\right)$$

$$=\frac{a_0}{2}+\sum_{l=1}^{k}(a_l\cos lx+b_l\sin lx),$$

所以

$$a_0=\frac{2}{\sqrt{2\pi}}x_0=\frac{1}{\pi}\int_0^{2\pi} f(x)\mathrm{d}x;$$

$$a_l=\frac{x_{2l-1}}{\sqrt{\pi}}=\frac{1}{\pi}\int_0^{2\pi} f(x)\cos lx\mathrm{d}x,\quad l=1,2,\cdots,k.$$

$$b_l=\frac{x_{2l}}{\sqrt{\pi}}=\frac{1}{\pi}\int_0^{2\pi} f(x)\sin lx\mathrm{d}x,\quad l=1,2,\cdots,k;$$

于是以上述 a_0, a_l, b_l 为系数的三角多项式

$$p(x)=\frac{a_0}{2}+\sum_{l=1}^{k}(a_l\cos lx+b_l\sin lx)$$

使 $\int_0^{2\pi}(f(x)-p(x))^2\mathrm{d}x$ 最小。称 $a_0,a_1,\cdots,a_k,b_1,\cdots,b_k$ 为 $f(x)$ 的 Fourier 系数。

习　题

1. 设W是 n 维欧氏空间V的非平凡子空间，且W对V的线性变换 σ 不变(即W是不变子空间)，求证：$W^{\perp}$ 是V的关于 σ' 的不变子空间。

2. 设 W_1 与 W_2 是欧氏空间V的两个有限维非平凡子空间，求证：

$$(W_1+W_2)^{\perp}=(W_1\cap W_2)^{\perp},$$
$$(W_1\cap W_2)^{\perp}=W_1^{\perp}+W_2^{\perp}$$

3．求矛盾方程组：

$$\begin{cases}3x_1-x_2=2\\ x_1+x_2=2\\ 2x+3x_2=4\end{cases}$$

的最小二乘解。

§5 酉空间概述

一、酉空间的基本性质

欧氏空间是实数域上带有内积(它是实函数)的线性空间，本节考虑基域是复数域，并带有内积(它是复函数)的线性空间，由于引进概念的想法及一些结论的推导绝大部分与欧氏空间相类似，故不需一一重复叙述与论证，只要把它与欧氏空间有区别的地方提出就可以了，例如内积的定义是下面的

定义 设 V 是复空间(复数域上的线性空间)，对 V 中任何两个向量 α、β，定义唯一的一个复值(泛函)函数，记为(α,β)，且满足：

(i) $(\alpha,\beta)=\overline{(\beta,\alpha)}$, $\quad\forall\alpha,\beta\in V$;

(ii) $(k\alpha,\beta)=k(\alpha,\beta)$, $\quad\forall\alpha,\beta\in V,k\in K$;

(iii) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$, $\forall\alpha,\beta,\gamma\in V$;

(iv) $(\alpha,\alpha)\geqslant0$，而等号成立的充要条件是，$\alpha=\theta$。则称$(\alpha,\beta)$为**内积**，称带有内积$(\alpha,\beta)$的复空间为**酉空间**。

内积的第(i)条之所以如此定义，乃是为了保证第(iv)条的(α,α)是一实数，这样才能谈到$(\alpha,\alpha)\geqslant0$。另一方面，这样定义内积，当 α、β 是实向量时，就把欧氏空间的内积作为这个内积的一个特例，因而酉空间是欧氏空间的推广。

关于酉空间的内积的初等性质中，有一点与欧氏空间不同，即

$$(\alpha,k\beta)=\overline{(k\beta,\alpha)}=\overline{k(\beta,\alpha)}=\bar{k}\,\overline{(\beta,\alpha)}=\bar{k}(\alpha,\beta).$$

凡在欧氏空间成立的一些基本性质在酉空间也成立，且其证

法也基本上相同，但对酉变换σ(即$(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta)$的酉空间中的线性变换)的如下结论的证法却不相同，即

定理 σ是n维酉空间的酉变换的充要条件是，对任何$\alpha\in V$，恒有

$$(\sigma(\alpha),\sigma(\alpha))=(\alpha,\alpha). \tag{1}$$

【证明】 此定理的证明与欧氏空间中的证明有所不同，故证明如下：

必要性是显然的。因为只要在定义$(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta)$中取$\beta=\alpha$即可。

充分性。因为对任何α、β，恒有(1)式，故

$$(\sigma(\alpha+\beta),\sigma(\alpha+\beta))=(\alpha+\beta,\alpha+\beta), \tag{2}$$

$$(\sigma(\alpha+\sqrt{-1}\beta),\sigma(\alpha+\sqrt{-1}\beta))=(\alpha+\sqrt{-1}\beta,\alpha+\sqrt{-1}\beta). \tag{3}$$

由(2)式并应用(1)式，显然可算得

$$(\sigma(\alpha),\sigma(\beta))+(\sigma(\beta),\sigma(\alpha))=(\alpha,\beta)+(\beta,\alpha). \tag{4}$$

同理，由(3)式与(1)式可算得

$$-(\sigma(\alpha),\sigma(\beta))+(\sigma(\beta),\sigma(\alpha))=-(\alpha,\beta)+(\beta,\alpha), \tag{5}$$

由(4)与(5)即得

$$(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta),$$

故σ是酉变换。 证毕。

选 做 题

1．设δ是欧氏空间V的非零向量，求证：对任何$\alpha\in V$，必有

$$\alpha_{L(\delta)}=\frac{(\alpha,\delta)}{\|\delta\|^2}\delta.$$

2．设$\alpha_1,\alpha_2,\cdots,\alpha_s$是欧氏空间$V$的正交向量组，

(i) 证明下面的 Bessel-Parsseval 不等式：

$$\|\alpha\|^2\geqslant\sum_{i=1}^{s}\|\alpha_{L(\alpha_i)}\|^2;$$

(ii) 当$\alpha_1,\alpha_2,\cdots,\alpha_s$取作标准正交向量组$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s$时，求证：

$$\|\alpha\|^2\geqslant\sum_{i=1}^{s}(\alpha,\varepsilon_i)^2 \tag{Δ}$$

(即向量的长度不小于它在各个"正交坐标轴"$L(e_i)$ 上的正射影的长度的平方和)；

(iii) 当V是有限维空间时，(Δ)式是等式(此时变为推广的商高定理)的充要条件是，$e_1, e_2, \cdots, e_s$是V的标准正交基(因而$d(V)=s$)。

3. 设α_W是α在欧氏空间V的非平凡子空间W上的正射影，证明：$\sigma: \alpha \to \alpha_W$是自共轭($\sigma=\sigma'$)、幂等($\sigma^2=\sigma$)线性变换，称$\sigma$为正射影变换。

4. 求证：欧氏空间V的自共轭、幂等线性变换必是正射影变换。